

AD _____

AWARD NUMBER DAMD17-98-1-8600

TITLE: A Micro-simulation Model of the Benefits and Costs of
Prostate Cancer Screening and Treatment

PRINCIPAL INVESTIGATOR: David O. Meltzer, MD, Ph.D.

CONTRACTING ORGANIZATION: University of Chicago
Chicago, Illinois 60637

REPORT DATE: July 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20000511 101

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE July 1999	3. REPORT TYPE AND DATES COVERED Annual (1 Jul 98 - 30 Jun 99)	
4. TITLE AND SUBTITLE A Micro-simulation Model of the Benefits and Costs of Prostate Cancer Screening and Treatment			5. FUNDING NUMBERS DAMD17-98-1-8600	
6. AUTHOR(S) David O. Meltzer, MD, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of Chicago Chicago, Illinois 60637			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Although prostate cancer has immense human and financial costs, there is great uncertainty about how to best screen for and treat prostate cancer. This project uses data from existing studies of prostate cancer to develop a decision model to estimate the costs and benefits of screening and treatment for prostate cancer. Using new methods that account for the fact that prostate cancers may progress quickly or slowly and that doctors may not be certain of the actual pathologic stage of cancer when making treatment decisions, we are developing more accurate estimates of the benefits and costs of screening and treatment than have previously been available. In assessing the benefits of screening and treatment we consider the effects on both the length and quality of life. We are also using our model to look ahead to developing technologies to try to judge which ones may offer the biggest benefits to patients to help guide research dollars to the most promising advances. So far, we have been able to complete on time all the tasks we detailed on our original proposal. We expect our analysis to make a substantial contribution to prostate cancer research.				
14. SUBJECT TERMS Prostate Cancer			15. NUMBER OF PAGES 85	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

✓ ____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

David M. ... 7/20/99
PI - Signature Date

(4) Table of Contents

	Page
(1) Front Cover.....	1
(2) Standard Form.....	2
(3) Foreword.....	3
(4) Table of Contents.....	4
(5) Introduction.....	5
(6) Body.....	5
(7) Key Research Accomplishments.....	13
(8) Reportable Outcomes.....	13
(9) Conclusions.....	14
(10) References.....	15
(11) Appendices.....	21

(5) Introduction:

Although prostate cancer has immense human and financial costs, there is great uncertainty about how to best screen for and treat prostate cancer. This project uses data from existing studies of prostate cancer to develop a decision model to estimate the costs and benefits of screening and treatment for prostate cancer. Using new methods that account for the fact that prostate cancers may progress quickly or slowly and that doctors may not be certain of the actual pathologic stage of cancer when making treatment decisions, we are developing more accurate estimates of the benefits and costs of screening and treatment than have previously been available. In assessing the benefits of screening and treatment we consider the effects on both the length and quality of life. We are also using our model to look ahead to developing technologies to try to judge which ones may offer the biggest benefits to patients to help guide research dollars to the most promising advances.

(6) Body:

Our progress to date has closely followed the work specified in the **Statement of Work** section of our original proposal. We have been able to accomplish our specific goals with little or no delay. In general, our model is comprised of two main features. The first is a model of the natural history of prostate cancer. This involves a state vector that allows us to carry states of prostate cancer symptoms and stages from year to year and a matrix of transition probabilities that describe the progression of prostate cancer in the absence of screening or treatment. The second is a decision tree that allows us to model the likelihood that a man has prostate cancer detected in a year given the disease state that he carries from the beginning of the year (see Appendix 1). Much of our work to date has been devoted to determining the probabilities that are necessary for the decision tree and transition between states. One difference from our proposal concerns the time period from which the data for our estimates will come. In the original proposal, we had assumed that our estimates would reflect the prevalence and treatment of prostate cancer from the late 1980s and early 1990s. Since then, a number of new studies and data sources have been released that will enable us to derive our estimates from 1990s data. This will make our research more up-to-date and more relevant. We will outline our progress in relation to the **Statement of Work** point by point below.

Task 1. Review of model structure to reflect evolving innovation, Months 1-3.

a) Relevance given current knowledge of prostate biology (i.e. grade, p53, early metastases)

Prostate cancer has often been characterized as a slow-growing tumor. However, the natural history of prostate cancer varies tremendously based on tumor grade and stage.^{1,2} We have spent considerable time reviewing literature on the biology of prostate cancer to determine which aspects of its biology affect treatment decisions and outcomes. Such information has allowed us to determine what states are necessary for us to carry in the state vector from year to year, and the probabilities that a person transitions among states. Our comprehensive literature review has allowed us to reassess our original hypotheses concerning the natural history of prostate cancer, and hence revise some of our models. At the same time, our literature review has helped us clarify the strengths and limitations of current knowledge concerning the natural history, diagnosis, and treatment of prostate cancer.

One important development in prostate cancer research has been the identification of p53, a tumor suppressor gene that has been found to be mutated in a majority of cancer cell lines.³ However, the frequency of p53 mutation in men with prostate cancer is only about 15%, a much lower incidence than its mutation in other cancers.^{4,5} Preliminary data suggest that the p53

mutation is associated with prostate cancer progression to hormone refractory state and is more frequent in metastatic lesions.⁶ Although the p53 mutation appears to be associated with more advanced prostate cancer, its utilization in daily patients' management is premature at this point and further studies are needed to confirm its prognostic value. We therefore have not incorporated information concerning p53 mutations as markers for early metastases into our model. However, our model is designed to be flexible enough to include new developments concerning p53 into our decision analysis if its mutation is likely to be adopted as a prognostic tool.

On the other hand, the value of stage and grade in determining the outcomes of prostate cancer is well documented. Therefore, we have included prostate cancer grade in our state vector that we carry from year to year.

Task 1. Review of model structure to reflect evolving innovation, Months 1-3.

b) Relevance given current screening and treatment technology (i.e. PSA density, % free PSA)

The most relevant technology for detecting prostate cancer continues to be prostate-specific antigen (PSA) testing. In addition to its role in detection, prostate specific antigen is increasingly being used to clinically stage prostate cancer tumors. This is because serum PSA level has been shown to be of prognostic value for patients with prostate cancer. Thus the outcome of any treatment modality has been shown to worsen as the clinical stage, PSA, and Gleason⁷ grade increase. Although the AJCC staging system for prostate cancer does not take serum PSA into consideration and the grade is only partially integrated into the staging system, many recent publications group their patient populations into 3 prognostic groups (favorable, intermediate, and unfavorable) according to all the important prognostic factors, including PSA, Gleason, and clinical stage. The favorable group includes patients with T1-2, PSA < 10, and Gleason < 7. The intermediate group includes patients with only one prognostic factor of a higher value than stated for the favorable group (either PSA > 10, or Gleason > 6, or clinical stage > T2). The unfavorable group includes patients with 2 or more poor prognostic values. Given the diagnostic and prognostic value of PSA tests, we have incorporated PSA level detectability as a state in the state vector of the model so that this may be studied in detail.

Considering specific types of PSA tests, it remains to be seen whether PSA density will offer a clinical benefit compared to standard PSA tests.⁸ Likewise, the data is still too limited to determine whether free PSA is clinically useful.^{9,10,11} As such, we have not formally incorporated such measurement techniques into the model, but have programmed our model to permit their inclusion later if new research determines that they have diagnostic or prognostic value.

Task 1. Review of model structure to reflect evolving innovation, Months 1-3.

c) Relevance for potential future technologies (i.e. PCR, Indium-111 labeled antibody to detect metastases, gene therapy)

We have not found that any of the new therapies listed above (PCR, Indium-111 labeled antibody to detect metastases, gene therapy) have come into widespread use as diagnostic or treatment tools for men with prostate cancer. However, we have designed our model to be flexible so that we can allow for the effects of new diagnostic and therapies on prostate cancer detection, treatment, and outcomes. Thus, if any of these new technologies develop to the point that they are introduced into standard practice, we have ensured that we can easily incorporate their effects on prostate cancer detection, treatment, and cost-effectiveness practices into our model. In addition, we have done a substantial amount of theoretical work developing a theoretical framework with which to assess the potential value of information about these technologies, as we discuss below.

Task 2. Programming of model to reflect revised structure, Months 1-6
a) Revision of decision tree if needed

We have made substantial improvements in our decision tree compared to the one that we submitted for review in the original proposal. Our changes have mostly been designed to allow us to make our model more accurate, but have also allowed us to simplify the model in some places to improve our computational efficiency. For example, in the original grant, we submitted a decision tree with multiple branches tracing the likelihood that a person would have symptoms associated with prostate cancer, receive diagnostic tests, receive specific treatments for prostate cancer, and have specific complications or outcomes due to their treatments. The decision tree was completely symmetric for ease of interpretation. However, because there were near zero probabilities associated with a significant number of branches of the decision tree, this was very inefficient computationally. For example, it is virtually impossible that someone would receive a transurethral resection of the prostate (TURP) given that they do not have benign prostatic hyperplasia (BPH) and its associated symptoms. We therefore eliminated TURP as a possible mode of discovering prostate cancer among asymptomatic men. While this reduced the symmetry of our model, it also greatly reduced the number of possible branches in the decision tree. This helps with programming and presentation of the model. Appendix 1 depicts the revised decision tree.

Task 2. Programming of model to reflect revised structure, Months 1-6
b) Revision of natural history model if needed

As mentioned above, we have rethought some of our assumptions concerning the natural history of prostate cancer. In the original model, for example, we assumed that a given percentage of men would have BPH symptoms each year. We modeled the probability that men would have symptoms as independent of whether they had symptoms the previous year. However, this approach would have overestimated the cumulative number of men who have had BPH over time because men who do not have BPH in one year are less likely to have it the next year. Hence, the probability of developing BPH and associated symptoms should be conditional on whether or not symptoms were present in the previous year. To address this, we have created a large vector with 740 states that we carry from year to year. The states that are included in this vector include combinations of BPH symptoms (two states), cancer symptoms (two states), rectal and PSA detectability (6 states), having had a TURP (2 states), true stage (5 states), and true grade (3 states). We assume that the probability of moving into each of the states is conditional on which state the person was in the previous year. One point to clarify is that rectal and PSA detectability states indicate whether a person would have detectable prostate cancer if they had each test. It does not indicate whether or not a person actually has had a PSA or digital rectal exam test. Also, we carry the true stage and grade through the model as opposed to the clinical stage and grade that might be diagnosed after a biopsy, digital rectal exam (DRE), or prostate-specific antigen (PSA) test. At clinical staging nodes of the model, there is a certain likelihood that the true (pathologic) grade and stage will be misclassified as a different clinical grade and stage. The true grades and stage of prostate cancer are only often determined after radical prostatectomies in which the entire prostate is removed and examined.

Task 2. Programming of model to reflect revised structure, Months 1-6
c) Programming of costs
d) Programming of benefits and quality of life adjustments

We have chosen to discuss the programming of costs along with the programming of benefit and quality of life adjustments together since both sections affect our model in similar ways. To date, we have not changed the basic method of incorporating costs, benefits, and quality

of life assumptions into our program. As a result, we have not needed to do much reprogramming of the basic cost equations in our model.

Most of the complex programming to incorporate costs into our model will come when we complete programming sensitivity analyses over the next six months. At that point, we will need to determine how varying our assumptions about treatment patterns and screening test sensitivity alters the cost effectiveness of tests and treatments, with techniques described in task 4 below.

Task 3. Review of literature to estimate model parameters, Months 4-18

a) Identification of parameters of interest

In revising our decision tree model, we have identified the parameters of importance. Appendix 1 shows the decision tree. Each node of the decision tree represents at least one probability parameter that needs to be estimated.

Task 3. Review of literature to estimate model parameters, Months 4-18

b) Collection of articles relevant for parameters of interest

c) Review and analysis of articles

d) Preparation of documentation of literature review

We have been working on parts a, b, and c of task 3 simultaneously. For each parameter of interest, we have completed a comprehensive literature review to obtain desired parameter estimates. Where the parameter estimates are not readily available from the literature review, we have conducted our own data analyses. We have been formatting the analyses for presentation in a comprehensive document describing how we have derived our estimates. Appendix 2 provides an example of how we are documenting the methods used to derive each parameter estimate. Below are brief summaries of several of the key parameters used in the model and our progress in estimating them. Some of these parameters are used in single nodes in the decision tree. Others are used at several nodes in the tree.

Parameter: Prevalence of prostate cancer across stages

To derive the prevalence of prostate cancer across stages, we reviewed the existing literature that addresses the issue of prostate cancer prevalence by pathological examination of the prostate glands of men who either died of causes other than prostate cancer (autopsy series) or of men who had their prostate glands surgically removed for disease other than prostate cancer (pelvic exenteration, cystoprostatectomy). It is well documented that meticulous pathological examination of the prostate gland (5-mm slices, whole mount sections) is superior to routine random pathologic evaluation (prostate cancer prevalence almost doubles when whole mount sections examination is done in comparison to simple random examination). We used exclusively studies that utilized whole mount sections of the prostate gland. All studies indicate an increase in the pathological prevalence of prostate cancer with age.^{12,13,14,15,16,17,18,19,20} We have also tried to find similar studies to determine the likelihood that the stage distribution worsens with increasing age, but found very little evidence at all.²¹ Therefore, we will have to develop our natural history model largely in the absence of such data.

Parameter: Prevalence of prostate cancer across grades

To derive the prevalence of prostate cancer across grades we relied on studies that use the Gleason grading system to report the grade of prostate cancer. In the Gleason grading system the primary (the most dominant) and secondary pattern (the second most dominant pattern) are given a grade from 1 to 5. Gleason score is the sum of both digits, which range from 2 to 10. In our model, tumors with Gleason Score of 2 to 4 are considered well differentiated, 5 to 7 moderately differentiated, and 8 to 10 poorly differentiated. In our model we used the prevalence of the true

grade of prostate cancer seen in the final radical prostatectomy specimen. Since there might be a discrepancy in Gleason Score between prostate biopsy grade and ultimate prostatectomy grade given that biopsy sample just a small portion of the prostate, we based our estimates on grade noted in studies of radical prostatectomies.²²

One problem with the Gleason score is that it may be an imprecise method of grading tumors since it is based on the overall composition of different grades in a tumor, not just the worst grade. Hence, an individual with a tumor composed of both moderately and poorly differentiated prostate cancer would receive a different Gleason score than an individual who only had poorly differentiated grades of prostate cancer, even though the fact that both had poorly differentiated tumors that might cause them to have similar outcomes regardless of the presence of better differentiated grade within the tumor. However, given that the Gleason score is so widely used in the literature, we have decided to use Gleason scoring as the way that we classify well, moderately, and poorly differentiated tumors.

Parameter: Prevalence of benign prostatic hyperplasia (BPH) and cancer symptoms

We wish to estimate here the probability that BPH or cancer symptoms would prompt a man to have a digital rectal exam or PSA test. For BPH symptoms, we performed a meta-analysis of epidemiological data to define the prevalence of symptomatic BPH. Deriving the probability that someone had cancer symptoms was not quite so straightforward since not everyone with cancer has symptoms. Therefore, we used a series of equations at differing ages that relate the underlying probability that someone has symptoms given that they have prostate cancer to the conditional odds ratio of having cancer given that symptoms occur. The specific methods are detailed in Appendix 2. Solving these equations gave us the necessary probabilities for the likelihood of having symptoms given that a man has cancer.

Parameter: Rectal and PSA sensitivity and specificity

We used a number of articles to determine sensitivity and specificity of PSA and rectal examination.^{23,24,25,26,27,28,29,30,31,32, 33,34}

Some of these studies are classic screening studies that use multiple modalities of screening to determine sensitivity and specificity. Others (e.g. Chodak et al.) are population-based cohort studies. From these, we can determine the likelihood that someone who has a positive PSA test or digital rectal exam has prostate cancer. We can combine the population prevalence of prostate cancer with the likelihood that someone has prostate cancer given that they have a positive DRE or PSA test to determine the sensitivity and specificity of these tests.

Parameter: Sensitivity/specificity of biopsy tests

We assume that all men who have positive DRE or PSA tests have a biopsy to determine if they have prostate cancer. The grade of prostate cancer may be misclassified by the biopsy. For example, since a single tumor may consist of multiple grades, the final grade diagnosis may be contingent on which part of the tumor is sampled.

In order to determine the sensitivity and specificity of biopsies, both overall and by grade, we constructed a meta-analysis of the reported literature to determine the sensitivity and specificity of prostate gland biopsy in detecting the true grade. Based on criteria defined by Bailar and Mosteller,³⁵ in addition to identifying all relevant studies that have been reported in the literature (by conducting a MEDLINE search and citations from related articles^{22,36,37,38,39,40,41,42,43,44, 45,46,47,48}), the criteria to include studies in the our meta-analysis were the following:

- 1) Original data must be presented. Studies that included or overlapped with previously published data, and review papers were carefully evaluated but were not in the meta-analysis.
- 2) The studies must quantify grade on a common scale. Grading should be reported in three tier system (well-differentiated, moderately differentiated, and poorly differentiated), using the Gleason score.

- 3) The data must be reported in enough detail to allow full analysis and replication. The actual number of patients with their biopsy and true grade (rather than just percentage) must be reported.
- 4) The pathologist(s) should be unaware of the true grade when reviewing the biopsy grade.
- 5) The type of biopsy should be specified and involve multiple cores or ultrasound guidance (XX)
- 6) The pathologic examination of the radical prostatectomy specimen performed to compare biopsy results with true grade and stage must be thorough. The ideal method of evaluating the resected prostate gland is to perform a whole mount sections with thin (5 mm) slices.
- 7) There should be no patient selection bias, i.e., a randomly selected population with masking of the investigators to variables being investigated, or consecutive patients that were seen within a specific period of time.
- 8) There should be a comprehensive staging work up of the patients. This includes detailed history and physical examination, DRE, blood tests (Alkaline phosphatase, PSA, acid phosphatase), bone scan, CT-scan.
- 9) There should be no hormonal blockade prior to prostatectomy.

The first 3 criteria are considered essential for a study to be included in the meta-analysis. Criteria 4 to 9 were considered important and but we accepted studies that did not meet some of these criteria. Almost all of the studies are single institution retrospective reviews that compare the grade in biopsy with that of radical prostatectomy specimens. Few studies compared consecutive patients within a specified period of time. We excluded one study⁴⁹ because it included only patients whose biopsy revealed moderately differentiated tumors.

The probabilities resulting from this meta-analysis are ready to be incorporated into the model. A formal presentation of the results will be completed shortly.

Parameter: Probability of having a transurethral resection of the prostate (TURP)

Prostate cancer may be detected after examination of the removed prostate sample once a person has a TURP to relieve symptoms caused by BPH. Therefore, it is important to know the rate at which TURPs are performed. We are currently undertaking a literature review to determine the likelihood that someone with BPH related symptoms has a TURP. If we cannot obtain reliable estimates from the literature, we will use maximum likelihood estimation to derive them by using our model and national data on TURP rates by age.

Parameter: Probability of detection via TURP by stage

In order to capture all men who have prostate cancer detected, it is important to know the likelihood that prostate cancer is detected following a TURP. A review of the literature found that among men with no clinical evidence of prostate cancer who undergo TURP for urinary obstructive symptoms, prostate cancer is found in about 17.5%.^{50,51,52,53,54,55}

To determine the sensitivity of TURP in prostate cancer diagnosis, prostate cancer prevalence from our compiled prevalence data was utilized. The expected prevalence of occult prostate cancer in men of age 70 and older is about 40-70%. Given that TURP detect cancer in about 17.5%, its sensitivity is about 33%. As such, TURP has a sensitivity of 33% in detecting occult prostate cancer and about 17.5% of patients who undergo TURP will have cancer found in the resected tissue.

Parameter: Metastatic work-up positive

Part of the staging of prostate cancer includes determining if metastases have spread to distant parts of the body. We are currently reviewing the literature to determine the proportion of men with different stages of prostate cancer who have metastases as evidenced by staging examinations or progression following prostatectomy. If we cannot adequately determine this parameter from the literature, we can use maximum likelihood estimation to estimate it based on the model.

Parameter: Treatment

Knowing the treatments given stage of prostate cancer is important to understanding processes of care and outcomes among those with prostate cancer. We have conducted a study using the 1995 Surveillance, Epidemiology, and End Results (SEER) data from the National Cancer Institute to determine standards of care given the clinical stage of prostate cancer. In the past, many studies had published rates of prostate cancer treatment based on a combination of true stage and clinical stage. Since physicians base treatment decisions on the clinical stage of prostate cancer -- indeed, the true stage is only known after treatment with a prostatectomy -- we felt it was important to conduct our own study to determine treatment rates based on clinical stage.

A copy of a paper detailing our study that is currently under review has been added in Appendix 3. The general findings of our study indicate that treatment rates by most accurate stage differ substantially from rates by clinical stage. For example, surgery rates for stages B, C, and D are 33%, 68%, and 7% based on most accurate stage but 24%, 8%, and 1% based on clinical stage. Treatment rates by clinical stage vary dramatically by age. Over 60% of men age 50-54 with clinical stage A and B choose surgery versus 20% of men age 70-74, and 1% after age 80. Radiation rates for clinical stage A and B rise with age up to age 70-74, before declining in older patients. About half of men with clinical stage C up to age 80 choose radiation, but almost 40% of men age 50-54 choose surgery.

Our study includes only treatments of watchful waiting, radiation, and radical prostatectomy. Our model and sensitivity analyses are equipped to allow us to include any other treatments that we later determine have either widespread use or potentially significant effects on outcomes.

Parameter: Complication rates given treatments

The complications resulting from treatments for prostate cancer can have a major impact on quality of life and hence the cost-effectiveness of particular treatments. In order to determine the rate of complications associated with prostate cancer treatments, a Medline search was carried out to identify prospective and retrospective studies that address the complication rates. We identified the main quality of life issues in men treated for localized prostate cancer to be urinary continence, urethral stricture, sexual potency, and bowel function. There is a general agreement in the literature that radical prostatectomy is associated with higher rates of treatment-induced urinary incontinence and impotence than radiotherapy, while radiotherapy results in more bowel dysfunction. Nonetheless, the majority of treated patients are satisfied with their choice of therapy and they would choose the same treatment again.^{56,57,58} However, the satisfaction rates appear to decline over time in patients who do develop treatment complications.⁵⁹

Major sources of information came from the Prostate Outcome Research Team report to the Office of Technology Assessment of the U. S. Congress in 1995,⁶⁰ and other related studies.^{61,62,63,64} In the end, our best estimates of the complication rates were:

	Radical <u>Prostatectomy</u>	External Beam <u>Radiation Therapy</u>
Treatment associated mortality	1.1%	0.2%
Incontinence	33.4%	7.3%
Bowel injury	4.0%	13.7%
Urethral stricture requiring long-term treatment	12.4%	4.5%
Impotence	85.0%	40.0%

One consideration worthy of discussion is that the introduction of nerve-sparing radical prostatectomy has changed the rate of impotence after surgery for prostate cancer. While the impotence rate after non-nerve sparing radical prostatectomy is almost 90-100%, the rate is lower

with nerve sparing radical prostatectomy, especially in younger patients. We will use the sensitivity analyses of our study to determine how increased use of nerve sparing radical prostatectomy will affect complication rates.

Parameter: Cure rates and long-term mortality following treatment

Eliminating prostate cancer is an important goal of many treatments. Our model assumes that radical prostatectomies among those with stage A or stage B cancers is curative. We also assume that stage C or D prostate cancer cannot be cured. We are currently conducting literature reviews to determine the long-term mortality associated with radiation treatment for prostate cancer.

Parameter: Transition rates across states

Knowing how men transition across states in our state vector is an important component of our research. For example, many men with stage B cancer will progress to stage C or stage D cancer if they are not treated. These transitions will occur at the end of the decision tree model for those men who are not found to have prostate cancer, or have non-terminal prostate cancer. Once a man is actually found to have terminal prostate cancer, he no longer cycles through the program. Only men at risk of developing worsened prostate cancer cycle from year to year. As described in our initial proposal, these will be estimated over the next year using maximum likelihood techniques.

Task 4: Sensitivity Analysis: Months 6-18

- a) *All parameters in model*
- b) *One and multi-way sensitivity analysis*

We currently are programming the sensitivity analyses for the model. Most of the high and low estimates of parameters have already been estimated as described in the parameter derivations above.

Task 4: Sensitivity Analysis: Months 6-18

- c) *Assess effects on:*
 - 1) *Costs*
 - 2) *Effectiveness*
 - 3) *Cost-effectiveness*

We have done a great deal of time investigating the theoretical basis of cost-effectiveness analysis.^{65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81} Dr. Meltzer has recently written a manuscript entitled, "Addressing Uncertainty in Cost-Effectiveness Analysis: Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research." A copy of the paper has been included in Appendix 4. This paper examines the purposes for which sensitivity analysis is performed in medical cost-effectiveness analysis and the implications of an expected utility maximization model for the methods to perform such analyses. When sensitivity analysis is done to help guide decision making in the presence of uncertainty, expected utility maximization implies that the optimal decision is described by the ratio of the mean cost divided by the mean benefit. When sensitivity analysis is done to make decisions for subgroups with identifiable differences in costs or benefits, the standard is the same, only using the appropriate mean costs and mean benefits for each group. When sensitivity analysis is done to set priorities for the collection of additional information, the expected utility maximization framework suggests that value of information calculations can be done in the same context to assess the value of research. We will use such formulations in our own prostate cancer model.

In addition, we have been doing a great deal of research on determining patient utilities for different states with prostate cancer and treatment complications for the cost-effectiveness analysis. In general, utility assessment is a formal measurement of patients' preferences for their health states. It reflects the strength of the preference or the degree of abhorrence for the potential outcome of interest. Utilities are usually expressed numerically from 1.0 (utility for perfect health) to 0.0 (utility for death). The most common method of utility assessment is the time trade-off method. In time trade-off, patients are asked to estimate the survival time they are willing to sacrifice in order to retain perfect health. In our model, as in other Markov analyses, utility values are associated with each state of health. Studies that address utility in prostate cancer patients were identified by Medline search. We have currently reviewed numerous studies that use time-trade off questions to assess patient utilities.^{82,83,84,85,86,87,88}

(7) Key Research Accomplishments

- Created more accurate model concerning the natural history of prostate cancer and a more accurate decision tree concerning the treatment of prostate cancer.
- Identified relevant natural history, treatment, and outcome parameters for inclusion in our natural history and decision tree models.
- Determined the model parameter probabilities using literature reviews, meta-analyses, and primary data analysis.
- Submitted to a peer-reviewed journal a paper describing prostate cancer treatment rates across clinical stages.
- Submitted to a peer-reviewed journal a paper describing the theoretical basis for addressing uncertainty in cost-effectiveness analysis

(8) Reportable Outcomes

Articles:

- *Meltzer D, Eggleston BL, Abdalla I. Patterns of Prostate Cancer Treatment by Clinical Stage and Age in the United States. Under review, 1999.
- *Meltzer D. Addressing Uncertainty in Cost-Effectiveness Analysis: Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research. Under review, 1999.
- *Meltzer D, Johannesson M. Inconsistencies in the 'Societal Perspective' on Costs of the Panel on Cost-Effectiveness in Health and Medicine. In press, *Medical Decision Making*.
- *Meltzer D, Johannesson M. On the Role of Theory in Cost-Effectiveness Analysis: A Response To Garber, Russell, and Weinstein. In press, *Medical Decision Making*.

Published Abstracts

- *Meltzer D, Hazen G, Johannesson M, Abdalla I, Chang R, Elstein A, Schwartz A. Effect of Future Costs on the Cost-Effectiveness of Life Extension and Quality of Life Improvement Among the Elderly. *Medical Decision Making* 18: 475, October 1998.
- *Meltzer D, Polonsky T. Do Quality-Adjusted Life Years Reflect Patient Preferences? Validation Using Revealed Preference for Intensive Treatment of Insulin-Dependent Diabetes Mellitus.

Medical Decision Making 18: 459, October 1998.

Presentations:

"Implications of Expected Utility Maximization for Assessing Cost-Effectiveness under Uncertainty: Methods for Sensitivity Analysis and Implications for the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research". National Bureau of Economic Research Summer Institute, Boston, MA, July 1998.

"Effect of Future Costs on the Cost-Effectiveness of Life Extension and Quality of Life Improvement Among the Elderly", Society for Medical Decision Making Annual Meeting, Plenary Address, Boston, MA, October 1998.

"Accounting for Future Costs from a Societal Perspective in Cost-Effectiveness Analysis": Plenary Address, International Health Economics Association Annual Meeting, Rotterdam, Netherlands, June 1999.

"Measuring the Burden of Illness". Invited Presentation to Dr. Harold Varmus and panel of experts convened at NIH in Response to IOM Report on Scientific Priorities at NIH, June 1999.

(9) Conclusions

We have made substantial progress in refining our models and comprehensively reviewing the relevant literature concerning the natural history of prostate cancer, its treatment, and the associated outcomes. We have made substantial progress towards determining the following parameters of our model and are completing formal documentation of these findings:

- The prevalence of prostate cancer in men over the age of 50 by stage and age.
- The prevalence of benign prostatic hyperplasia and prostate cancer related symptoms.
- The likelihood that a man has prostate cancer that may be detected by a prostate screening antigen (PSA) test or by a digital rectal exam (DRE).
- The likelihood that a biopsy after a positive DRE or PSA test will correctly or incorrectly classify the stage and grade of prostate cancer.
- The likelihood that prostate cancer will be found after a transurethral resection of the prostate (TURP).
- How prostate cancer is treated across stage and age.
- The likelihood that a man will have a complication after a given treatment.
- The utility associated with living in specific health states after developing complications due to curative or palliative treatment for prostate cancer.

We have also furthered the field of cost-effectiveness research by developing sound theoretical models on the most efficient and appropriate method for conducting sensitivity analyses.

So what do our findings mean for the future of the field?

- Prostate cancer screening and treatment remain controversial because existing research has not yet provided a convincing, comprehensive assessment of the value of intervention for prostate cancer. Our research is highly innovative in that we are building on past research to build a comprehensive model that will detail the natural history of prostate cancer, and the most cost-effective screening policies and treatments. Our model will be flexible enough to allow for the evaluation of any new therapies or screening tests that are developed in the near future.
- Our manuscript detailing how prostate cancer is treated across clinical stage and age is the first to report national statistics on treatment decisions in a clinically relevant way. In the past authors had often used the "best information" available from a combination

of pathologic and clinical stage to detail how prostate cancer is treated according to stage. This gave a biased view of the way that physicians treated cancer upon presentation. Our paper allows health professionals and patients to see how prostate cancer is actually treated. This can assist patient and physician decision-making. Our findings also provide relevant information concerning standard practice patterns to researchers and policymakers who are researching and debating appropriate treatment for various stages of prostate cancer.

- Our manuscript developing theoretical underpinnings of cost-effectiveness research provides relevant methods for cost-effectiveness research on almost any subject. If the practices detailed by the paper are widely adopted, inconsistencies in current cost-effectiveness analyses will be reduced. This could make cost-effectiveness analyses more comparable across studies.

Our research is by no means complete. However, we are well on our way to finishing the project on time. So far, we have been able to complete on time all the tasks we detailed on our proposal. We continue to expect our analysis to make a substantial contribution to prostate cancer research.

(10) References

- ¹ Chodak GW, Thisted RA, Gerber GS, Johansson JE, Adofsson J, Jones GW, Chisholm GD, Moskovitz B, Livne PM, Warner J. Results of conservative management of clinically localized prostrate cancer. *NEJM* 1994; 330:242-248.
- ² Johansson JE, Amadio HO, Johansson S, Bergstrom R, Adami HO. Fifteen-year survival in prostate cancer, a prospective, population-based study in Sweden. *JAMA* 1997; 277:467-471
- ³ deVere White RW, Gumerlock PH, Chi S-G, Meyers FJ. p53 tumor suppressor gene abnormalities are frequent in human prostate tissues. *Journal of Urology* 1993; 149:376A.
- ⁴ Bookstein R, MacGrogan D, Hilsenbeck SG, Sharkey F, Allred DC. p53 is mutated in a subset of advanced-stage prostate cancers. *Cancer Research* 1993; 53(14):3369-3373.
- ⁵ Visakorpi T, Kallioniemi OP, Heikkinen A, Koivula T, Isola J. Small subgroup of aggressive, highly proliferative prostatic carcinomas defined by p53 accumulation. *Journal of the National Cancer Institute* 1992; 84:883-887.
- ⁶ Navone NM, Troncoso P, Pisters LL, Goodrow TL, Palmer JL, Nichols WW, von Eschenbach AC, Conti CJ. p53 protein accumulation and gene mutation in the progression of human prostate carcinoma. *Journal of the National Cancer Institute* 1993; 85:1657-1669
- ⁷ Gleason DF. Histologic grade, clinical stage, and patient age in prostate cancer. *NCI Monographs* 1988; 7:15-18.
- ⁸ Mettlin C, Littrup PJ, Kane RA, Murphy GP, Lee F, Chesley A, Badalament R, Mostofi FK. Relative sensitivity and specificity of serum prostate specific antigen (PSA) level compared with age-referenced PSA, PSA density, and PSA change: Data from the American Cancer Society National Prostate Cancer Detection Project. *Cancer* 1994; 74:1615-1620.
- ⁹ Prestigiacomo AF, Lilja H, Pettersson K, Wolfert RL, Stamey TA. A comparison of the free fraction of serum prostate specific antigen in men with benign and cancerous prostates: The best case scenario. *Journal of Urology* 1996; 156:350-354.

- ¹⁰ Pannek J, Rittenhouse HG, Chan DW, Epstein JI, Walsh PC, Partin AW. The use of percent free prostate specific antigen for staging clinically localized prostate cancer. *Journal of Urology* 1998; 159:1238-1242.
- ¹¹ Catalona WJ, Partin AW, Slawin KM, Brawer MK, Flanigan RC, Patel A, Richie JP, deKernion JB, Walsh PC, Scardino PT, Lange PH, Subong EN, Parson RE, Gasior GH, Loveland KG, Southwick PC. Use of the percentage of free prostate-specific antigen to enhance differentiation of prostate cancer from benign prostatic disease: A prospective multicenter trial. *JAMA* 1998; 279:1542-1547.
- ¹² Baron E, Angrist A. Incidence of occult adenocarcinoma of the prostate. *Archives of Pathology* 1941; 32:787-793.
- ¹³ Edwards CN, Steinthorsson E, Nicholson D. An autopsy study of latent prostate cancer. *Cancer* 1953; 6:531-554.
- ¹⁴ Franks LM. Latent carcinoma of the prostate. *Journal of Pathology and Bacteriology* 1954; 68:603-616.
- ¹⁵ Halpert B, Schmalhorst WR. Carcinoma of the prostate in patients 70 to 79 years old. *Cancer* 1966; 19:695-698.
- ¹⁶ Scott R, Mutchnik DL, Laskowski TZ, Schmalhorst WR. Carcinoma of the prostate in elderly men: Incidence, growth, characteristics, and clinical significance. *Journal of Urology* 1969; 101:602-607.
- ¹⁷ Lundberg S, Berge T. Prostatic carcinoma: An autopsy study. *Scandinavian Journal of Urology and Nephrology* 1970; 4:93-97.
- ¹⁸ Sakr WA, Haas GP, Cassin BF, Pontes JE, Crissman JD. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *Journal of Urology* 1993; 150:379-385.
- ¹⁹ Sakr WA, Gifnon DJ, Crissman JD, Heilbrun LK, Cassin BJ, Pontes JJ, Haas GP. High grade prostatic intraepithelial neoplasia (HGPIN) and prostatic adenocarcinoma between the ages of 20-69: An autopsy of 249 cases. *In Vivo* 1994; 8:439-443.
- ²⁰ Guileyardo JM, Johnson WD, Welsh RA, Akazaki K, Correa P. Prevalence of latent prostate carcinoma in two U.S. populations. *Journal of the National Cancer Institute* 1980; 65:311-316.
- ²¹ McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA. Patterns of progression in prostate cancer. *Lancet* 1986; 1:60-63.
- ²² Steinberg DM, Sauvageot J, Piantadosi S, Epstein JI. Correlation of prostate needle biopsy and radical prostatectomy Gleason grade in academic and community settings. *American Journal of Surgical Pathology* 1997; 21:566-576.
- ²³ Scaletsky R, Koch MO, Eckstein CW, Bicknell SL, Gray GF, Smith JA. Tumor volume and stage in carcinoma of the prostate detected by elevations in prostate specific antigen. *Journal of Urology* 1994; 152:129-131.
- ²⁴ Reissigl A, Horninger W, Fink K, Klocker H, Bartsch G. Prostate carcinoma screening in the county of Tyrol, Austria: Experience and results. *Cancer* 1997; 80:1818-1829.
- ²⁵ Rietbergen JBW, Kranse R, Kirkels WJ, De Koning HJ, Schroder FH. Evaluation of prostate-specific antigen, digital rectal examination and transrectal ultrasonography in population-based

screening for prostate cancer: Improving the efficiency of early detection. *British Journal of Urology* 1997; 79:57-63.

²⁶ Mettlin CJ, Murphy GP, Ho R, Menck HR. The national cancer data base report on longitudinal observations on prostate cancer. *Cancer* 1996; 77:2162-2166.

²⁷ Gann PH, Hennekens Ch, Stampfer MJ. A prospective evaluation of plasma prostate-specific antigen for detection of prostatic cancer. *JAMA* 1995; 273:289-294.

²⁸ Flanigan RC, Catalona WJ, Richie JP, Ahmann FR, Huson MA, Scardino PT, DeKernion JB, Ratliff TL, Kavoussi LR, Dalkin BL, Waters WB, MacFarlane MT, Southwick PC. Accuracy of digital rectal examination and transrectal ultrasonography in localizing prostate cancer. *Journal of Urology* 1994; 152:1506-1509.

²⁹ Richie JP, Ratliff TL, Catalona WJ, Kavoussi LR, Ahmann FR, Dalkin BL, Hudson MA, Waters WB, Scardino PT, MacFarlane MT, Flanigan RC, Southwick PC, DeKernion JB. *Urology* 1993; 42:365-374.

³⁰ Catalona WJ, Richie JP, Ahmann PR, Hudson MA, Scardino PT, Flanigan RC, Dekernion JB, Ratliffe TL, Kavoussi LR, Dalkin BL, Waters WB, MacFarlane MT, Southwick PC. Comparison of digital rectal examination and serum prostate specific antigen in the early detection of prostate cancer: Results of a multicenter clinical trial of 6,630 men. *Journal of Urology* 1994; 151:1283-1290.

³¹ Brawer MK, Laboratory studies for the detection of carcinoma of the prostate. *Urologic Clinics of North America* 1990; 17:759-768.

³² Catalona WJ, Smith DS, Ratliff TL, Dodds KM, Coplen DE, Yuan JJJ, Petros JA, Adriole GL. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *New England Journal of Medicine* 1991; 324:1156-1161.

³³ Cooner WH, Mosley BR, Rutherford CL, Beard JH, Pond HS, Terry WJ, Igel TC, Kidd DD. Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. *Journal of Urology* 1990; 143:1146-1154.

³⁴ Chodak GW, Keller P, Schoenberg HW. Assessment of screening for prostate cancer using the digital rectal examination. *Journal of Urology* 1989; 1136-1138.

³⁵ Bailar JC, Mosteller F. *Medical uses of statistics.* (Boston: NEJM books; 1992)

³⁶ Mills SE, Fowler JE. Gleason histologic grading of prostatic carcinoma: Correlations between biopsy and prostatectomy specimens. *Cancer* 1985; 57:346-349.

³⁷ Thickman D, Speers WC, Philpott PJ, Shapiro H. Effect of the number of core biopsies of the prostate on predicting Gleason score of prostate cancer. *Journal of Urology* 1996; 156:110-113.

³⁸ Djavan B, Kadesky K, Klopukh B, Marberger M, Roehrborn CG. Gleason scores from prostate biopsies obtained with 18-gauge biopsy needles poorly predict Gleason scores of radical prostatectomy specimens. *European Urology* 1998; 33:261-270.

³⁹ Spires SE, Cibull ML, Wood DP, Miller S, Spires SM, Banks ER. Gleason histologic grading in prostatic carcinoma: Correlation of 18-gauge core biopsy with prostatectomy. *Archives of Pathology and Laboratory Medicine* 1994; 118:705-708.

⁴⁰ Humphrey PA, Baty J, Keetch D. Relationship between serum prostate specific antigen, needle biopsy findings, and histopathologic features of prostatic carcinoma in radical prostatectomy tissues. *Cancer* 1995; 75:1842-1849.

- ⁴¹ Fernandes ET, Sundaram CP, Long R, Soltani M, Ercole CJ. Biopsy Gleason score: How does it correlate with the final pathological diagnosis in prostate cancer? *British Journal of Urology* 1997; 79:615-617.
- ⁴² Bostwick DG. Gleason grading of prostatic needle biopsies: Correlation with grade in 316 matched prostatectomies. *American Journal of Surgical Pathology* 1994; 18:796-803.
- ⁴³ Babaian RJ, Grunow WA. Reliability of Gleason grading system in comparing prostate biopsies with total prostatectomy specimens. *Urology* 1985; 25:564-567.
- ⁴⁴ Catalona WJ, Stein AJ, Fair WR. Grading errors in prostatic needle biopsies: Relation to the accuracy of tumor grade in predicting pelvic lymph node metastases. *Journal of Urology* 1982; 127:919-922.
- ⁴⁵ Lange PH, Narayan P. Understaging and undergrading of prostate cancer: Argument for postoperative radiation as adjuvant therapy. *Urology* 1983; 21:113-118.
- ⁴⁶ Kojima M, Troncoso P, Babaian RJ. Use of prostate-specific antigen and tumor volume in predicting needle biopsy grading error. *Urology* 1995; 45:807-812.
- ⁴⁷ Garnett JE, Oyasu R, Grayhack JT. The accuracy of diagnostic biopsy specimens in predicting tumor grades by Gleason's classification of radical prostatectomy specimens. *Journal of Urology* 1984; 131:690-693.
- ⁴⁸ Catalona WJ, Stein AJ, Fair WR. Grading errors in prostatic needle biopsies: Relation to the accuracy of tumor grade in predicting pelvic lymph node metastases. *Journal of Urology* 1982; 127:919-922.
- ⁴⁹ Carlson GD, Calvanese CB, Kahane H, Epstein JI. Accuracy of biopsy Gleason scores from a large uropathology laboratory: Use of a diagnostic protocol to minimize observer variability. *Urology* 1998; 51:525-529.
- ⁵⁰ Murphy WM, Dean PJ, Brasfield JA, Tatum L. Incidental carcinoma of the prostate. How much sampling is adequate?. *American Journal of Surgical Pathology* 10:170-174, 1986.
- ⁵¹ Ornstein DK, Rao GS, Smith DS, Andriole GL. The impact of systematic prostate biopsy on prostate cancer incidence in men with symptomatic benign prostatic hyperplasia undergoing transurethral resection of the prostate. *Journal of Urology* 1997; 157:880-883.
- ⁵² Bostwick DG, Cooner WH, Denis L, Jones GW, Scardino PT, Murphy GP. The association of benign prostatic hyperplasia and cancer of the prostate. *Cancer* 1992; 70:291-301.
- ⁵³ Moore GH, Lawshe B, Murphy J. Diagnosis of adenocarcinoma in transurethral resectates of the prostate gland. *American Journal of Surgical Pathology* 1986; 10:165-169.
- ⁵⁴ Lu-Yao GL, Barry MJ, Chang C, Wasson JH, Wennberg JE. Transurethral resection of the prostate among Medicare beneficiaries in the United States: time trends and outcomes. *Urology* 1994; 44:692-699.
- ⁵⁵ McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *American Journal of Surgical Pathology* 1988; 12:897-906.
- ⁵⁶ Herr HW. Quality of life in prostate cancer patients. *Ca: a Cancer Journal for Clinicians* 1997; 47:207-217.

- ⁵⁷ Fowler FJ Jr, Barry MJ, Lu-Yao G, Wasson JH, Bin L. Outcomes of external-beam radiation therapy for prostate cancer: a study of Medicare beneficiaries in three surveillance, epidemiology, and end results areas. *Journal of Clinical Oncology* 1996; 14:2258-2265.
- ⁵⁸ Wasson JH, Cushman CC, Bruskewitz RC, Littenberg B, Mulley AG, Wennberg JE. A structured literature review of treatment for localized prostate cancer. *Archives of Family Medicine* 1993; 2:487-493.
- ⁵⁹ Herr HW. Quality of life of incontinent men after radical prostatectomy. *Journal of Urology* 1994; 151:652-654.
- ⁶⁰ US Congress, Office of Technology Assessment. Costs and Effectiveness of Prostate Cancer Screening in Elderly Men, OTA-PH-H-145. (Washington, DC: US Government Printing Office, May, 1995).
- ⁶¹ Wasson JH, Cushman CC, Bruskewitz RC, Littenberg B, Mulley AG, Wennberg JE. A structured literature review of treatment for localized prostate cancer. *Archives of Family Medicine* 1993;2:487-493.
- ⁶² Herr HW. Quality of life in prostate cancer patients. *CA: a Cancer Journal for Clinicians* 1997; 47:207-217.
- ⁶³ Feldman HA, Goldstein I, Hatzichristou DG, Krane RJ, McKinlay JB. Impotence and its medical and psychosocial correlates: results of the Massachusetts Male Aging Study. *Journal of Urology* 1994; 151:54-61.
- ⁶⁴ Laumann EO, Paik A, Rosen RC. Sexual dysfunction in the United States: prevalence and predictors. *JAMA* 1999; 281:537-544.
- ⁶⁵ Barry MJ, Fowler FJ Jr, Mulley AG Jr, Henderson JV Jr, Wennberg JE. Patient reactions to a program designed to facilitate patient participation in treatment decisions for benign prostatic hyperplasia. *Medical Care* 1995; 33:771-782.
- ⁶⁶ Cutler D, Richardson E. Measuring the Health of the U.S. Population. *Brookings Papers: Microeconomics* 1997:217-271.
- ⁶⁷ Erickson P, Wilson R, Shannon I. Years of Healthy Life. *Healthy People 2000 Statistical Notes*. Centers for Disease Control and Prevention/National Center for Health Statistics 7, April 1995.
- ⁶⁸ Gold MR, Siegel JR, Russel LB, Weinstein MC, eds. *Cost-Effectiveness in Health and Medicine*. (New York, NY: Oxford University Press, 1996).
- ⁶⁹ Institute of Medicine. *Scientific Opportunities and Public Needs: Improving Priority Setting and Public Input at the National Institutes of Health*. (Washington, DC: National Academy Press, 1998).
- ⁷⁰ Johannesson M, Meltzer D, O'Connor R. Incorporating Future Costs in Medical Cost-Effectiveness Analysis: Implications for the Cost-Effectiveness of the Treatment of Hypertension. *Medical Decision Making* 1997; 17:382-389.
- ⁷¹ Johannesson M, Meltzer D. Some reflections on cost-effectiveness analysis. *Health Economics* 1998; 7:1-7.
- ⁷² Manton KG, Corder L, Stallard E. Chronic trends in elderly United States populations: 1982-1994. *Proceedings of the National Academy of Sciences* 1997; 94:2593-2598.

- ⁷³ Meltzer D. Accounting for future costs in medical cost-effectiveness analysis. *Journal of Health Economics* 1997; 16:33-64.
- ⁷⁴ Meltzer D, Johannesson M. On the Role of Theory in Cost-Effectiveness Analysis: A Response to Garber, Russell and Weinstein. *Medical Decision Making*. In Press.
- ⁷⁵ Meltzer D. Implications of Expected Utility Maximization for Assessing Cost-Effectiveness under Uncertainty: Methods for Sensitivity Analysis and Implications for the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research. *Journal of Health Economics*; under review.
- ⁷⁶ Meltzer D, Johannesson M. Inconsistencies in the 'Societal Perspective' on Costs of the Panel on Cost-Effectiveness in Health and Medicine. *Medical Decision Making*; in press.
- ⁷⁷ Meltzer D, Polonsky T, Tobian J. Does Quality-Adjusted Life Expectancy Predict Patient Preferences for Treatment in IDDM: Validating Quality of Life Measures through Revealed Preference. In preparation.
- ⁷⁸ Murphy KM, Topel RH. The Economic Value of Medical Research. Unpublished manuscript, University of Chicago, 1998.
- ⁷⁹ Nord E. Methods for Quality Adjustment of Life Years. *Social Science and Medicine* 1992; 34:559-569.
- ⁸⁰ Tengs, T. Planning for Priorities Project, No. 2, Progressive Policy Institute, July, 1998.
- ⁸¹ Weinstein M, Stason WB. *Hypertension: A Policy Perspective*. (Cambridge: Harvard University Press; 1976).
- ⁸² Fleming C, Wasson JH, Albertsen PC, Barry MJ, Wennberg JE. A decision analysis of alternative treatment strategies for clinically localized prostate cancer. *JAMA* 1993; 269:2650-2658.
- ⁸³ Bennett CL, Chapman G, Elstein AS, Knight SJ, Nadler RB, Sharifi R, Kuzel T. A comparison of perspectives on prostate cancer: analysis of utility assessments of patients and physicians. *European Urology* 1997; 32(suppl 3):86-88.
- ⁸⁴ Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA* 1998; 280:975-980.
- ⁸⁵ Cowen ME, Miles BJ, Cahill DF, Giesler RB, Beck JR, Kattan MW. The danger of applying group-level utilities in decision analyses of the treatment of localized prostate cancer in individual patients. *Medical Decision Making* 1998; 18:376-380.
- ⁸⁶ Kattan MW, Cowen ME, Miles BJ. A decision analysis for treatment of clinically localized prostate cancer. *Journal of General Internal Medicine* 1997; 12:299-305.
- ⁸⁷ Cantor SB, Spann SJ, Volk RJ, Cardenas MP, Warren MM. Prostate cancer screening: a decision analysis. *Journal of Family Practice* 1995; 41:33-41.
- ⁸⁸ Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS. Screening for prostate cancer: a decision analytic view. *JAMA* 1994; 272:773-780.

(11) Appendices

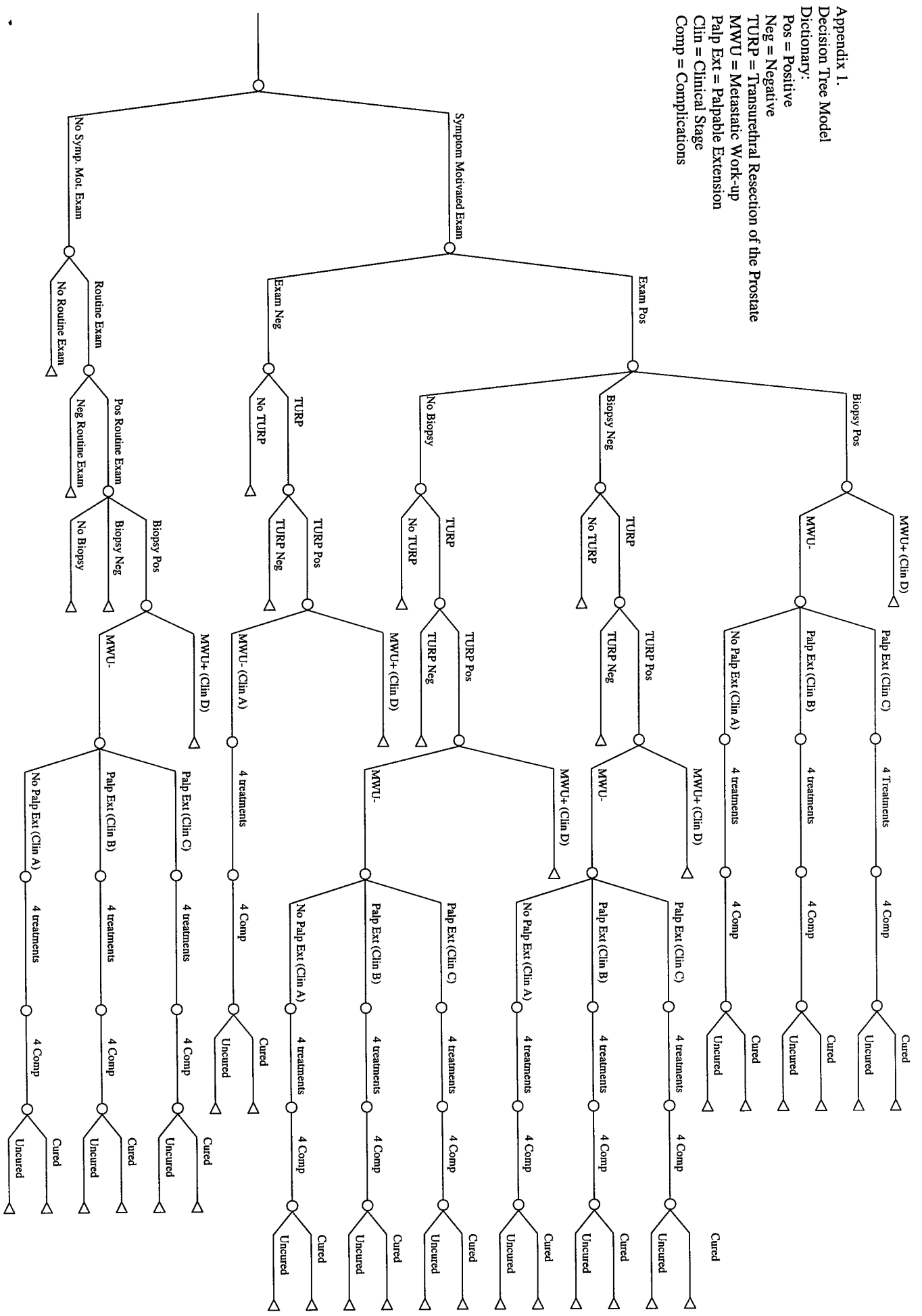
Appendix 1. Decision Tree Model.

Appendix 2. Cancer symptoms documentation. This provides an example of the documentation being generated for all parameters in the model.

Appendix 3. Meltzer D, Eggleston BL, Abdalla I. Patterns of Prostate Cancer Treatment by Clinical Stage and Age in the United States. Under review, 1999.

Appendix 4. Meltzer D. Addressing Uncertainty in Cost-Effectiveness Analysis: Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research. Under review, 1999.

Appendix 1.
 Decision Tree Model
 Dictionary:
 Pos = Positive
 Neg = Negative
 TURP = Transurethral Resection of the Prostate
 MWU = Metastatic Work-up
 Palp Ext = Palpable Extension
 Clin = Clinical Stage
 Comp = Complications



Parameter Name: Probability of symptoms given cancer

Description: This section describes the probability that someone will have symptoms similar to BPH given that they have prostate cancer.

Derivation Technique: We determined from the literature the odds ratio of having cancer given symptoms, the prevalence of cancer, and the prevalence of BPH. We used this to mathematically solve for the probability of having symptoms given cancer.

Introduction:

This section describes the probability that someone has symptoms given prostate cancer. It is important to know this information so that we can determine the proportion of men who have symptoms that lead them to get exams. The condition of having symptomatic prostate cancer is a state that is carried from year to year so that we do not overestimate the prevalence of symptomatic prostate cancer in the population. In this study, we assume that tumors only cause symptoms if they are large enough.

Derivation:

The estimates are obtained by using published information concerning the odds ratio of having cancer given symptoms and the prevalence of large tumor prostate cancer and BPH in the population. This information is used in the following way:

Let C represent having prostate cancer, S represent having symptoms, B represent having BPH, and P_x represent the probability of having a condition where x is replaced by B(BPH), C (cancer), CB (prostate cancer and BPH), S (symptoms), -S (no symptoms).

$$\text{Odds Ratio for Cancer(OR)} = \frac{P(C|S)}{P(C|-S)} = \frac{P(C|S) / P(S)}{P(C|-S) / (1 - P(S))} = \frac{P(S|C)P(C) / P(S)}{(1 - P(S|C))P(C) / (1 - P(S))} = \frac{P(S|C) / P(S)}{(1 - P(S|C)) / (1 - P(S))}$$

Rearranging this:

$$(1 - P(S|C)) * OR = P(S|C) * \left(\frac{1 - P(S)}{P(S)} \right)$$

$$OR = P(S|C) * \left[OR + \left(\frac{1 - P(S)}{P(S)} \right) \right]$$

$$P(S|C) = \frac{OR}{OR + \left(\frac{1 - P(S)}{P(S)} \right)}$$

$$OR = \frac{P(C|S)}{P(C|-S)} = \frac{P(CS) / P(S)}{P(C-S) / (1 - P(S))} = \frac{(P_{CB} + P(S|C)P_C) / (P_{CB} + P_B + P(S|C)P_C)}{(1 - P(S|C))P_C / [(1 - P(S|C))P_C + (1 - P_C - P_B - P_{CB})]}$$

$$(P_C - P(S|C)P_C)(P_{CB} + P_B + P(S|C)P_C)OR = (P_{CB} + P(S|C)P_C)[P_C + (P(S|C)P_C + (1 - P_C - P_B - P_{CB}))]$$

Let $X = P(S|C)P_C$, then

$$(P_C - X)(P_{CB} + P_B + X)OR = (P_{CB} + X)[P_C + X + (1 - P_C - P_B - P_{CB})]$$

$$[-X^2 + (P_C - P_{CB} - P_B)X + (P_C P_{CB} + P_C P_B)]OR = X^2 + [P_C + (1 - P_C - P_B - P_{CB}) + P_{CB}]X + P_{CB}P_C + P_{CB}(1 - P_C - P_B - P_{CB})$$

Let $a = (1 - OR)$,

$$b = [P_C + (1 - P_C - P_B - P_{CB}) + P_{CB} - OR(P_C) + OR(P_{CB}) + OR(P_B)]$$

$$c = P_{CB}P_C + P_{CB}(1 - P_C - P_B - P_{CB}) - P_CP_{CB}OR - P_CP_BOR$$

$$X = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Solve for $P(S|C)$: $P(S|C) = X / P_C$

The parameters for these equations are calculated as follows:

1) Odds Ratio of cancer given symptoms:

Lower bound: 1.0

Low estimate: 1.5

Best estimate: 2.0

High estimate: 2.5

Upper bound: none

--Source: Adolfsson J, Helgason AR, Dickman P, Steineck G. Urinary and bowel symptoms in men with and without prostate cancer: Results from an observational study in the Stockholm area. *European Urology* 1998; 22:11-16.

2) Calculations of P_B , P_C , P_{BC}

First, we determined the prevalence of BPH from the literature:

Table 1: Prevalence of BPH in US:

Best Estimate:

Age	Incidence	Prevalence (drawn from incidence rates)
40-49	0.0094	0.1224
50-59	0.0313	0.3020
60-69	0.0513	0.5518
70-79	0.0591	0.7481
80+	0.0591	0.8630

--Source: Glynn et al., The development of benign prostatic hyperplasia among volunteers in the normative aging study. *American Journal of Epidemiology* 1985, 121(1):78-90

Next, we determined the prevalence of prostate cancer from the literature:

Table 2: Prevalence of Prostate Cancer in US, any stage:

Age	Prevalence
40-49	0.266667
50-59	0.247911
60-69	0.322751
70-79	0.395238
80+	0.469512

--Source: Meta-analysis of:

--Baron E, Angrist A. Incidence of occult adenocarcinoma of the prostate. *Archives of Pathology* 1941; 32:787-793.

--Edwards CN, Steinthorsson E, Nicholson D. An autopsy study of latent prostate cancer. *Cancer* 1953; 6:531-554.

---Franks LM. Latent carcinoma of the prostate. *Journal of Pathology and Bacteriology* 1954; 68:603-616.

--Halpert B, Schmalhorst WR. Carcinoma of the prostate in patients 70 to 79 years old. *Cancer* 1966; 19:695-698.

--Scott R, Mutchnik DL, Laskowski TZ, Schmalhorst WR. Carcinoma of the prostate in elderly men: Incidence, growth, characteristics, and clinical significance. *Journal of Urology* 1969; 101:602-607.

--Lundberg S, Berge T. Prostatic carcinoma: An autopsy study. *Scandinavian Journal of Urology and Nephrology* 1970; 4:93-97.

--Sakr WA, Haas GP, Cassin BF, Pontes JE, Crissman JD. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *Journal of Urology* 1993; 150:379-385.

--Guileyardo JM, Johnson WD, Welsh RA, Akazaki K, Correa P. Prevalence of latent prostate carcinoma in two U.S. populations. *Journal of the National Cancer Institute* 1980; 65:311-316.

Finally, we determined the proportion of prostate cancer that is large enough to actually cause symptoms.

Table 3: Proportion of prostate cancer tumors large enough to cause symptoms (>0.46 ml); weighted by race

Age	Black Men * (12.7% of pop)	White Men* (87.3% of pop)	Calculation of proportion:	Proportion
40-49	5/14	8/26	$=0.127*5/14+(1-0.127)*8/26$	0.313973
50-59	5/14	8/26	$=0.127*5/14+(1-0.127)*8/26$	0.313973
60-69	5/14	8/26	$=0.127*5/14+(1-0.127)*8/26$	0.313973
70-79	9/17	14/35	$=0.127*9/17+(1-0.127)*14/35$	0.416435
80+	9/17	14/35	$=0.127*9/17+(1-0.127)*14/35$	0.416435

*Population data comes from the US Bureau of the Census, Population Division, release PPL-91, "United States Population Estimates, by Age, Sex, Race, and Hispanic Origin, 1990 to 1997." On July 1, 1994, the estimated number of black men in the US totaled 15,491,000 (12.7% of the black/white total). The estimated number of white men in the US totaled 106,246,000 (87.3% of the black/white total).

--Source of tumor size proportion data: McNeal JE, Bostwick DG, Kindrachuk RA, Redwine EA, Freiha FS, Stamey TA. Patterns of progression in prostate cancer. *Lancet* 1986; 60-63.

3) Calculations of P(S|C):

We used the information on the prevalence of BPH, prostate cancer, and large tumors to calculate the parameters P_{CB} , P_B , and P_C as follows:

Table 4: Calculating P_{CB} , P_B , P_C

Age	Prevalence of BPH in the US (Table 1)	Prevalence of large prostate tumors in US (Table 2 * Table 3)	P_{CB} (Column 1 * Column 2)	P_B (column 1-column 3)	P_C (Column 2-Column 3)
40-49	0.1224	0.083726007	0.010251	0.11218	0.073475
50-59	0.3020	0.0778372	0.023504	0.27846	0.054333
60-69	0.5518	0.101335049	0.055921	0.49592	0.045414
70-79	0.7481	0.164591092	0.123123	0.62493	0.041468
80+	0.8630	0.195521449	0.168731	0.69425	0.026790

Table 5 breaks out the calculations for the quadratic equation where a, b, and c are the variables for the quadratic equation, whose derivation is outlined above, and X+ is the X that is calculated from the equation

$$X = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

and X- is the X calculated from the equation

$$X = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Table 5: Using the quadratic formula to solve for X

Age	X+	X-	a	b	c
40-49	0.036467	0.24668429	-4	1.132604	-0.03598
50-59	0.036157	0.45377075	-4	1.959709	-0.06563
60-69	0.034593	0.72446034	-4	3.036214	-0.10025
70-79	0.032855	0.94414726	-4	3.90801	-0.12408
80+	0.021005	1.10067156	-4	4.486705	-0.09248

In the end, the quadratic formula could be solved for two distinct values of X: X+ and X-. Only one value of X gave a logical answer, however. The solutions obtained from X+ gave answers between the values of 0 and 1. The solutions obtained from X- gave answers between the values of 3 and 81, which are not possible probability values. Hence, using the X obtained from X+, we see that the probability of having symptoms given cancer are as follows:

Table 6: Solving of P(S|C), Final Results:

OR=1.5		OR=2		OR=2.5	
Age	P(S C)	Age	P(S C)	Age	P(S C)
40-49	0.063839	40-49	0.125376	40-49	0.185548
50-59	0.138616	50-59	0.250977	50-59	0.345106
60-69	0.219203	60-69	0.367838	60-69	0.476356
70-79	0.261645	70-79	0.422457	70-79	0.532188
80+	0.27725	80+	0.438869	80+	0.545197

Patterns of Prostate Cancer Treatment by Clinical Stage and Age in the United States

David Meltzer M.D., Ph.D.¹⁻³, Brian Eggleston M.P.P.¹
Ibrahim Abdalla M.D.⁴

¹Section of General Internal Medicine, ²Department of Economics, ³Harris Graduate School of Public Policy Studies, and ⁴Department of Radiation and Cellular Oncology

University of Chicago

July 17, 1999

Word Count: 3406

This work was supported by the Department of Defense Prostate Cancer Research Program (DAMD 17-98-1-8600), National Institute of Aging (K12 AG00488-08), and the Robert Wood Johnson Generalist Physician Faculty Scholars Program (Dr.Meltzer).

Corresponding Author: David Meltzer M.D., Ph.D.

Address: 5841 S. Maryland Avenue MC 2007 Chicago, IL 60637.

Tel: (773) 702-0836 Fax: (773) 834-2238 E-mail: dmeltzer@medicine.bsd.uchicago.edu

Abstract:

Context: Prostate cancer treatment remains controversial, but patients and their physicians may benefit from knowing treatment decisions of patients of similar age and clinical stage. However, treatment patterns have often been reported using the "most accurate" stage incorporating pathologic stage if available from surgery and no previous study has reported treatment rates by age and clinical stage.

Objective: To examine prostate cancer treatment rates by age and clinical stage, and contrast these to rates by most accurate stage.

Design: A retrospective analysis.

Setting and Participants: All men diagnosed with prostate cancer in 1995 in the Surveillance, Epidemiology, and End Results cancer registry.

Main Outcome Measures: Treatment rates by watchful waiting, radiation, surgery, or combined radiation and surgery.

Results: Treatment rates by most accurate stage differ substantially from rates by clinical stage. For example, surgery rates for stages B, C, and D are 33%, 69%, and 7% by most accurate stage but 34%, 8%, and 1% by clinical stage. Treatment rates by clinical stage vary dramatically by age. Over 60% of men age 50-54 with clinical stage A and B choose surgery versus 21% age 70-74, and 1% after age 80. Radiation rates for clinical stage A and B rise with age up to age 70-74, before declining in older patients. About half of men with clinical stage C up to age 80 choose radiation, but almost 40% age 50-54 choose surgery.

Conclusions: Treatment rates by most accurate stage differ significantly from rates by clinical stage. Patients and physicians should be aware of variations in treatment decisions by age and clinical stage.

Word count: 250

Introduction

Prostate cancer is the most commonly detected form of cancer in American men.¹ However, there is considerable debate concerning the proper treatment of prostate cancer.^{2,3,4,5,6} Because the disease often progresses slowly, many older men with prostate cancer are unlikely to die from it.³ Because treatments can often result in complications such as incontinence and impotence,^{7,8} the impact of treatment on quality of life may be considerable.

In the absence of clear guidelines for the appropriate treatment of prostate cancer, patients and their physicians may benefit from knowledge of common treatment patterns. A number of studies have reported data concerning the rates at which different treatments for prostate cancer are used.^{9,10,11,12,13,14} Most of these studies simply report overall rates aggregated across all stages and ages.^{10,11,12,13,14} However, since treatment is generally tailored to patient age and tumor stage, these aggregate statistics may not provide much information for patients or their physicians concerning the choices of similarly situated patients. Moreover, these overall treatment rates may change over time based on changes in the distribution of tumors by stage and patient age, even in the absence of changes in treatment patterns conditional on age and stage.

Moreover, even when analyses stratify tumors by stage, they rarely report treatment rates based on the clinical stage at the time that treatment decisions were made. For example, Mettlin et al.⁹ report rates stratified by stage, but are not able to determine whether the stage is based on the clinical stage or the pathologic stage determined following prostatectomy. Since pathologic staging of prostate cancer often results in a different stage than clinical staging,¹⁵ reports of cancer treatment based on the best available information combining clinical and pathologic stage ("most accurate stage") do not represent how prostate cancer is actually being treated upon clinical presentation. For example, an increase in the number or fraction of clinical stage B cancers treated with surgery could appear as an increase in the number and fraction of stage C

cancers treated with surgery if cancers are classified by most accurate stage and many stage B tumors are found to be stage C based on the results of prostatectomy.

One study that does report treatment rates by clinical stage uses data from the American College of Surgeons and American Cancer Society National Cancer Data Base.^{16, 17} However, that study does not report treatment rates by age and other studies using the NCDP have reported stage based on the most accurate information combining clinical and pathologic staging.^{9, 13, 18} Analyses by clinical stage have not been possible using population-based data because cancer registries and other sources of population-based cancer data have generally reported stage based on the most accurate information combining clinical and pathologic staging. The Surveillance, Epidemiology, and End Results (SEER) cancer-incidence database began to record clinical and pathologic stages separately in 1995, allowing for the analysis of treatment patterns by clinical stage. This study uses the 1995 SEER data to calculate treatment rates by age according to the true clinical stage, and contrasts these to estimates of treatment rates by age according to stage based on most accurate stage, as has generally been reported in the past.

One study that does report treatment rates by age and clinical stage uses data from the American College of Surgeons and American Cancer Society National Cancer Data Base.¹⁹ However, this study does not report treatment rates by age and other studies using the NCDP. Unfortunately, this is not a nationally population-based sample of cancer cases, and therefore may not be representative of treatment rates in the population.²⁰ Such analyses by clinical stage have not been possible using population-based data either because cancer registries and other sources of population-based cancer data have generally reported stage based on the most accurate information combining clinical and pathologic staging, or because researchers have only reported stage based on a combination of clinical and pathologic stages.^{9, 13, 21} The Surveillance,

Epidemiology, and End Results (SEER) cancer-incidence database began to record clinical and pathologic stages separately in 1995, allowing for the analysis of treatment patterns by clinical stage. This study uses the 1995 SEER data to calculate treatment rates by age according to the true clinical stage, and contrasts these to estimates of treatment rates by age according to stage based on most accurate stage, as has generally been reported in the past.

Methods

This study uses data from 1995 Surveillance, Epidemiology, and End Results (SEER) program (SEER Cancer Incidence Public-Use Database, 1973-1995, August 1997 Submission). This was the first year that SEER reported prostate carcinoma stage based on the pre-prostatectomy and post-prostatectomy stages. The SEER program has been collecting data on cancer incidence and prevalence since 1973 from cancer registries that now cover approximately 14% of the U.S. population, and is believed to be reasonably representative of the U.S. population as whole (SEER web-site, www-seer.ims.nci.nih.gov). All men over the age of 50 who had a diagnosis of prostate cancer in 1995 were included in the study.

Using SEER*Stat software, we selected all cases of SEER-identified prostate cancer from 1995. We defined the clinical stage of tumors as follows: Stage A included SEER identified tumors defined of A or T1, stage B included tumors of B or T2, stage C included tumors defined as stage C or with local extension into an adjacent structure, and D included tumors with extensive local extension or distant metastases. In contrast, SEER also reports the "most accurate" stage of cancer according to any information available from prostatectomy. In these cases, any patient with localized disease without extension is classified automatically as stage B

since SEER does not recognize pathologic stage A as a discrete entity. In addition, we classified patients with positive lymph nodes as pathologic stage D.

We grouped treatment into four categories: watchful waiting, surgical treatment with prostatectomy, radiation therapy of any kind, or a combination of surgical treatment and radiation. In describing our results, we did not combine the men choosing combination therapy with those choosing surgery alone to calculate rates of surgery because the number choosing combination therapy is generally quite small and including those men would unnecessarily complicate the interpretation of our results. Since transurethral resection of the prostate (TURP) is a therapeutic treatment of symptoms rather than a curative treatment, we grouped those who had TURPs with those who did not have treatment.

We calculated the age-specific treatment rates for each of the four stages based on two types of staging. The first calculations used the previously available method under which only the most accurate stage was known. The second calculations used the pre-prostatectomy clinical stage.

Results

The sample had a total of 15,954 men over age 50 with a diagnosis of prostate cancer. Of these, 79 were excluded because they had cystoprostatectomies, and hence their prostate cancer might have been found before a traditional clinical staging had been performed. An additional 1,692 were excluded because clinical stage was specified only as A or B and 1,694 were excluded because clinical stage was specified as unknown. 560 men were excluded for a variety of reasons such as missing pathologic data for patients who had undergone prostatectomy, missing information on lymph node status, or a variety of miscellaneous data inconsistencies.

These excluded subjects did not differ substantially from included subjects in age, race, or pathologic stage in those with information available for comparison. Of the remaining 11,929 men with prostate cancer and well-defined clinical stage, 4,252 (36%) were clinical stage A, 6,103 (51%) were clinical stage B, 707 (6%) were clinical stage C, and 867 (7%) were clinical stage D (see figure 1). Breaking this group down by most accurate stage, 2,759 (23%) had stage A, 5,845 (49%) had stage B, 2,224 (19%) had stage C, and 1,097 (9%) had stage D. In addition, 4 individuals who had a clinical diagnosis of prostate cancer were found to have a pathologic noninvasive or intraepithelial (in situ) tumor.

Overall, 30% of those with prostate cancer are treated with surgery, 30% are treated with radiation, and 1% are treated with a combination of the two, while the remaining 39% receive expectant management, which may include hormonal therapy.

Figure 1 reports treatment by the stage based on most accurate stage, as has often been done in the past, while figure 2 reports treatment rates based on clinical stage, with no recoding based on pathologic stage following prostatectomy. Combining patients across all ages and combining patients by stage according to the most accurate stage, no stage A patients receive prostatectomies (since SEER does not define a pathologic stage A), while 33% of stage B and 69% of stage C patients receive prostatectomies. Analyzing the data by clinical stage, the fraction of patients receiving surgery for stages A and B is still slightly more than 30%, but is only 8% for clinical stage C and 1% of clinical stage D patients. Rates of radiation therapy by most accurate stage decrease from 37% and 36% for stages A and B to 13% for stage C but rates by clinical stage rise from 24% for stage A and 34% for stage B to 43% for stage C. About 15% of stage D patients receive radiation by either staging system.

Treatment rates vary substantially by age whether by most accurate stage (figure 3) or clinical stage (figure 4), with watchful waiting almost uniformly increasing by age, but especially after age 75. When using the most accurate stage, no men with stage A cancer have prostatectomies (by definition), but over 70% of the youngest men with stage B cancer, and about 80% for the youngest men with stage C cancer have prostatectomies. Using clinical stage, the rates of prostatectomy are about 67% for the youngest men with stage A and B cancer but decrease to 38% and 5% for the youngest men with stage C and stage D cancer. The overstatement of surgery rates using most accurate compared to clinical stage is even greater for men between age 55 and 74. Using most accurate stage substantially overstates the fraction of young men choosing radiation for stage A disease compared to the fraction by clinical stage and substantially understates the fraction of clinical stage C patients choosing radiation. Rates of combined surgery and radiation remain low at all ages using either staging system for all stages.

Focusing on the treatment profiles by age for clinical stage, surgery is chosen by more than half to two-thirds of men age 64 and younger with stage A and B cancer, and almost half of those age 65-69, and then falls rapidly after age 70. Radiation therapy is chosen for clinical stage A and B disease by only about 10-15% of men age 50-54, but rises to be the choice of 30-50% of patients in their 70s before declining substantially for patients age 80 and above. Among men with clinical stage C cancer, the rate of surgery falls from a high of 38% of those 50-54 years old to less than 1% of those over 80 years old. Other than the youngest group of men with clinical stage C disease, among whom radiation rates are decreased because of the fraction of men choosing surgery, the fraction of men choosing radiation drifts slowly down with age from 59% in men age 55-59 to 40% in men age 75-79, before declining rapidly to 11% for those above age 80. Watchful waiting was the predominant form of treatment among clinical Ds, rising

from 70% of those 50-54 years old to 89% of those over 80 years old. Radiation without surgery falls from 27% among those 55-59 years old to 11% of those over 80 years old. Surgery rates both with and without radiation are low in this group but appear substantially higher if assessed among D's classified by the best available information.

Table 1 displays how clinical stages are restaged based on pathologic information.

Because SEER does not recognize a pathologic stage A, all of those with clinical stage A cancer are reclassified after surgery. Nearly half of those with clinical stage B cancers and nearly a quarter of those with clinical stage C cancers are reclassified. Several men originally classified as having stage D cancer had their cancer stage downgraded after surgery. Among the 2,224 men with stage C cancer by most accurate stage, 1589 (71%) resulted from restaging following prostatectomy.

Discussion

This study reports how each clinical stage of prostate cancer was treated in 1995. This was made possible because SEER changed the method of recording prostate cancer grade such that the pre-prostatectomy and post-prostatectomy stages were both recorded. Our results indicate that surgery rates among patients with stage A prostate cancer are much higher than previously reported. In an analysis of treatments by most accurate stage, Mettlin et al.⁹ reported that fewer than 17% of patients with stage A disease received surgical treatment. In contrast, we found that over 30% of patients with clinical stage A disease received surgical treatment, with rates reaching almost 70% among the youngest men in the sample. Conversely, in the same study as above, Mettlin et al. found that over 52% of stage C cancer patients received surgery, while we found that only 8% of clinical stage C cancer patients received surgery. This

discrepancy is explained by the large fraction of cases staged as C by most accurate stage occur because of restaging as the result of prostatectomy.

Although Mettlin's 1993 analysis of the 1984 and 1990 American College of Surgeons data by clinical stage did not calculate stage-specific rates of treatment by age,¹⁶ our results across all ages are quite similar with the exception of a substantial increase in the rate of surgery among men with stage A prostate cancer from 6.5% in 1984 and 10.1% in 1990 to 33% in our data. Whether this represents a change in treatment patterns given stage or a shift towards the detection of localized prostate cancer in younger men who tend to elect more aggressive treatment, or other differences between the two populations cannot be determined without a reanalysis of the National Cancer Data Base Data.

We found that patterns of radiation treatment differ from previous reports as well, but not as dramatically as with surgery rates. For example, our study found that approximately 24% of patients with clinical stage A cancer and 34% with clinical stage B cancer received radiation treatment. Again, in the analysis by most accurate stage, Mettlin et al.⁹ found that at least 34% of stage A and 36% of stage B patients receive radiation treatment, and in the analysis by clinical stage, Mettlin et al.¹⁶ found radiation therapy rates of about 16% for stage A and 40% for stage B in both periods. The higher rate of treatment of stage A tumors by radiation in the SEER data by most accurate stage that is reported by Mettlin likely represents the recoding of many clinical A or B patients to higher stages based on the results of surgery, so that the fraction of final As and Bs receiving radiation is increased. Along with the results described above for surgery rates among clinical stage A tumors, the increase in fraction of clinical stage A tumors treated with radiation seems to point again to an increase in the aggressiveness with which stage A prostate cancer is being treated. The absence of a strong gradient in radiation rates by age suggests that

this increase is not due to a change in the age distribution of the patients found with stage A prostate cancer. This could be confirmed by reanalysis of the American College of Surgeons data used by Mettlin et al.¹⁶ that examined treatment rates by age. Among those with stage C or stage D cancer, where it is likely that radiation is used for palliative reasons, one also sees some increase in the use of radiation by clinical stage compared to the 1984 and 1990 data reported by Mettlin.¹⁶

Putting aside these changes over time, our results confirm that substantial restaging of prostate cancer occurs after surgery. Upstaging of clinical stage A cancer after prostatectomy is required by definition, since the SEER staging system does not recognize pathologic stage A cancer. Such upgrading has led to the underestimation of prostatectomy rates among clinical stage A patients in studies based on most accurate stage. Conversely, because so many cancers are upstaged to pathologic stage C cancers after surgery, past estimates of surgery rates among Cs based on the "most accurate" information, which included pathologic instead of clinical information when available, overestimated the rate of prostatectomy among patients with a clinical stage C.

Our results show the importance of calculating treatment rates based on both clinical stage and age. In contrast to prior studies that do not report rates by combinations of age and clinical stage, our results show surgery is by far the predominant method of treating stage A and stage B prostate cancer among younger men with prostate cancer. Among somewhat older patients with stage A or stage B prostate cancer, radiation is more frequently used. Given the known morbidities and uncertain benefits associated with treatment, it is not possible to say whether this is a desirable treatment pattern. Moreover, any treatment decision must also consider an individual patient's preferences and comorbidities. Nevertheless, knowledge of

these common treatment patterns may nevertheless be useful for prostate cancer patients and their physicians. In addition, the finding that most young men choose surgery or radiation provides some insight into the challenges involved in enrolling patients into randomized clinical trials of treatment versus watchful waiting such as the currently funded National Cancer Institute Prostate Cancer Intervention Versus Observation Trial (PIVOT).²²

Our results are also striking for the fact that 20% of younger men with clinical stage C cancer are also treated surgically, despite the decreased likelihood that surgery will be of benefit in these men and the associated risks of complications. With older data based on the "most accurate" information, some of this relationship could be explained as patients with clinically observed A or B disease who were then found to have stage C disease. Our findings suggest this is not the case and intensify the need to understand whether such surgery is justified.

One challenge in identifying treatment patterns is that we were not able to identify through SEER whether the lymph node involvement was found during the clinical evaluation prior to surgery, or as part of pelvic lymph node dissection performed immediately before a prostatectomy that was consequently aborted. This implies that some patients who elect surgical treatment but have surgery aborted due to the intraoperative discovery of positive lymph nodes, may have their treatment misclassified. To assess this, we also performed a sensitivity analysis under the assumption that clinical A, B, or Cs with lymph node involvement had correct clinical staging information, but had been intended to have prostatectomy even if such surgery was not noted in the SEER data. Our findings indicate that surgery rates increased only slightly for clinical As and Bs since fewer than 2% were reassigned. Surgery rates increased more for those with clinical stage C cancer since 8% of the sample was reassigned. The surgery rate with or without radiation for men with clinical stage C cancer rises from 9% to 17%.

Our results also indicate the need for cancer databases to retain information on both clinical and pathologic states. In addition to describing these clinical and pathologic stages, our results also demonstrate the importance of retaining information on the timing of diagnoses regarding lymph node involvement in prostate cancer. In the SEER database, it is unknown if the diagnosis was made during clinical staging, as might happen during a CT scan, or through lymphadenectomy immediately prior to a planned prostatectomy which was consequently aborted. In our results, alternate assumptions regarding means by which lymph node status was determined required changes in the assignment of the intended treatment among those with stage D cancer.

Knowing treatment rates for prostate cancer is important to understanding standard practice patterns for treating this disease. As more aggressive screening techniques detect prostate cancer at earlier stages, more men will be faced with the decision of how to treat their cancer. Current debates regarding appropriate treatment for prostate cancer have not been resolved. In the absence of such resolution, aggressive treatment among younger men with early clinical stages of the cancer is the norm. Reliable data on the patterns of treatment according to stage will be important as we attempt to determine how treatment patterns change in years ahead.

Table 1 Restaging of cancer after prostatectomy

		Pathologic Stage				Lymph Node Involvement*	Total
		in-situ	Stage B	Stage C	Stage D		
Clinical Stage	Stage A	1	838	605	0	40	1,484
	Stage B	3	1,129	983	1	128	2,244
	Stage C	0	8	51	0	64	123
	Stage D	0	2	1	3	69	75
Total		4	1977	1640	4	301	3926

		Pathologic Stage				Lymph Involvement*	Total
		in-situ	Stage B	Stage C	Stage D		
Clinical Stage	Stage A	0%	56%	41%	0%	3%	1493
	Stage B	0%	50%	44%	0%	6%	2235
	Stage C	0%	7%	41%	0%	52%	123
	Stage D	0%	3%	1%	4%	92%	75
Total		4	1977	1640	305	1.52432	3926

*In this group, no pathologic stage or a pathologic stage other than D was recorded in the SEER data, but lymph node involvement was noted. Hence, it is likely that many people had a prostatectomy aborted subsequent to a positive lymph node dissection immediately before surgery.

References:

1. Landis SH, Murray T, Bolden S, Wingo PA. Cancer statistics, 1998. *CA Cancer J Clin.* 1998;48(1):6-29.
2. Collins MM, Barry MJ. Controversies in prostate cancer screening. Analogies to the early lung cancer screening debate. *JAMA.* 1996 Dec 25;276(24):1976-9.
3. Johansson JE, Holmberg L, Johansson S, Bergstrom R, Adami HO. Fifteen-year survival in prostate cancer. A prospective, population-based study in Sweden. *JAMA.* 1997 Feb 12;277(6):467-71.
4. Albertsen PC. Early-stage prostate cancer. When is observation appropriate? *Hematol Oncol Clin North Am.* 1996 June 10;(3):611-25.
5. Kelly WK, Scher HI. Prostate cancer: how can a common disease be so controversial? *Curr Opin Oncol.* 1994 May 6;(3):318-31.
6. Fowler FJ Jr, Bin L, Collins MM, Roberts RG, Oesterling JE, Wasson JH, Barry MJ. Prostate cancer screening and beliefs about treatment efficacy: a national survey of primary care physicians and urologists. *Am J Med.* 1998 June;104(6):526-32.
7. Fowler FJ Jr, Barry MJ, Lu-Yao G, Wasson J, Roman A, Wennberg J. Effect of radical prostatectomy for prostate cancer on patient quality of life: results from a Medicare survey. *Urology.* 1995;45(6):1007-13.

8. Wasson JH, Cushman CC, Bruskewitz RC, Littenberg B, Mulley AG Jr, Wennberg JE. A structured literature review of treatment for localized prostate cancer. Prostate Disease Patient Outcome Research Team. *Arch Fam Med*. 1993;2(5):487-93.
9. Mettlin CJ, Murphy GP, Cunningham MP, Menck HR. National Cancer Data Base report on race, age, and region variations in prostate cancer treatment. *Cancer*. 1997;80(7):1261-1266.
10. Lu-Yao G, Greenberg RG. Changes in prostate cancer incidence and treatment in USA. *Lancet*. 1994;343:251-254.
11. Lu-Yao G, McLerran D, Wasson J, Wennberg JE. An assessment of radical prostatectomy: Time trends, geographic variation, and outcomes. *JAMA*. 1993;269(20):2633-2636.
12. Klabunde CN, Potosky AL, Harlan LC, Kramer BS. Trends and black/white differences in treatment for nonmetastatic prostate cancer. *Medical Care*. 1998;36(9):1337-1348.
13. Mettlin CJ, Murphy GP, Rosenthal DS, Menck HR. The National Cancer Data Base report on prostate carcinoma after the peak in incidence rates in the US. *Cancer*. 1998;83(8):1679-1684.
14. Harlan L, Brawley O, Pommerenke F, Wali P, Kramer B. Geographic, age, and racial variation in the treatment of local/regional carcinoma of the prostate. *Journal of Clinical Oncology*. 1005;13(1):93-100.

15. Partin AW, Kattan MW, et al: Combination of PSA, Clinical stage, and Gleason Score to predict pathological stage of localized prostate cancer. *JAMA*. 1997;277:1445-1451.
16. Mettlin C, Jones GW, Murphy GP. Trends in prostate cancer care in the United States, 1974-1990: Observations from the patient care evaluation studies of the American College of Surgeons Commission on Cancer. *CA Cancer J Clin*. 1993;43:83-91.
17. Steele GD, Winchester DP, Menck HR. The National Cancer Data Base: A mechanism for assessment of patient care. *Cancer*. 1994;73(2):499-504.
18. Mettlin CJ, Murphy GP, Ho R, Menck HR. The national cancer data base report on longitudinal observations on prostate cancer. *Cancer*. 1996;77(10):2162-2166.
19. Mettlin C, Jones GW, Murphy GP. Trends in prostate cancer care in the United States, 1974-1990: Observations from the patient care evaluation studies of the American College of Surgeons Commission on Cancer. *CA Cancer J Clin*. 1993;43:83-91.
20. Steele GD, Winchester DP, Menck HR. The National Cancer Data Base: A mechanism for assessment of patient care. *Cancer*. 1994;73(2):499-504.
21. Mettlin CJ, Murphy GP, Ho R, Menck HR. The national cancer data base report on longitudinal observations on prostate cancer. *Cancer*. 1996;77(10):2162-2166.

22. Wilt TJ, Brawer MK. The Prostate Cancer Intervention Versus Observation Trial (PIVOT).
Oncology. 1997;11(8):1133-9; discussion 1139-40, 1143.

Figure 1: Prostatectomy and radiation treatment rates for prostate cancer according to stage defined by most accurate information

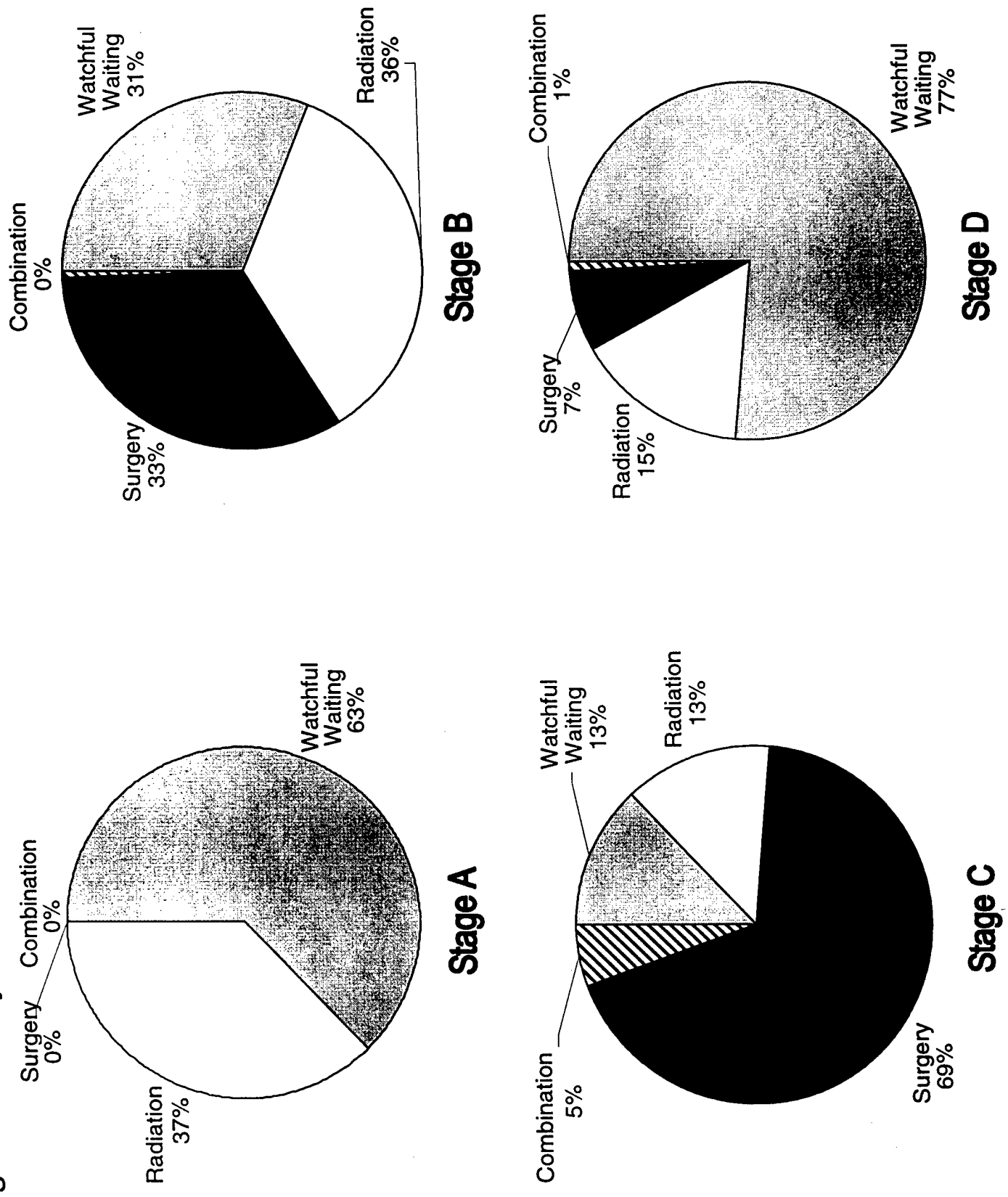


Figure 2: Prostatectomy and radiation treatment rates for prostate cancer according to stage defined by clinical diagnosis

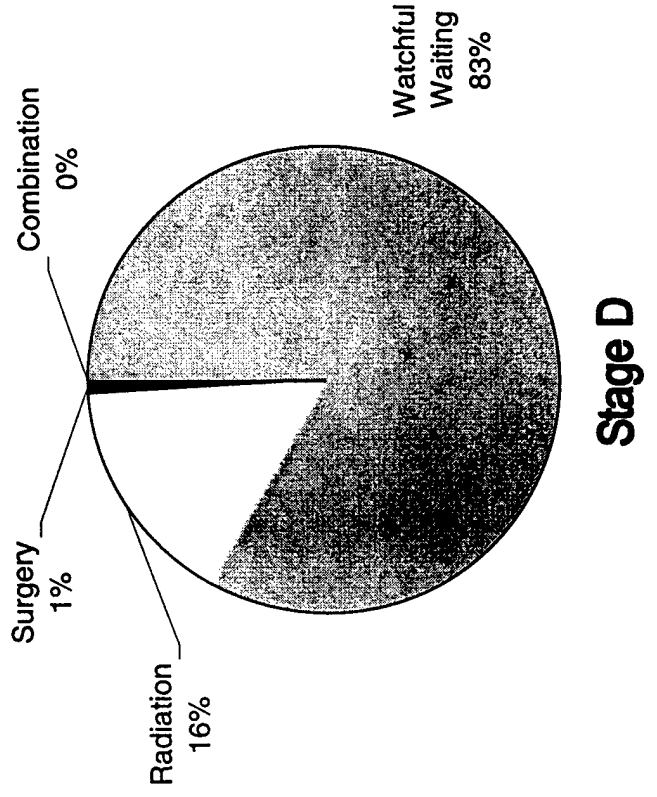
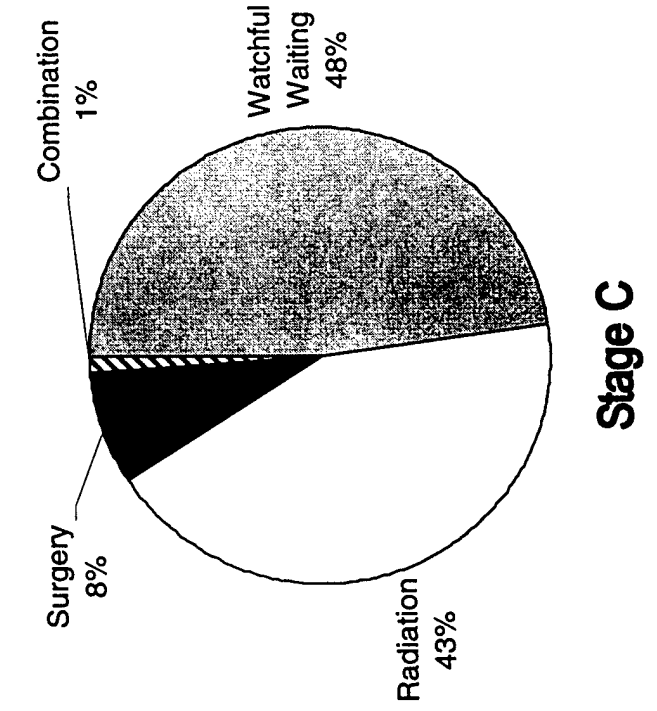
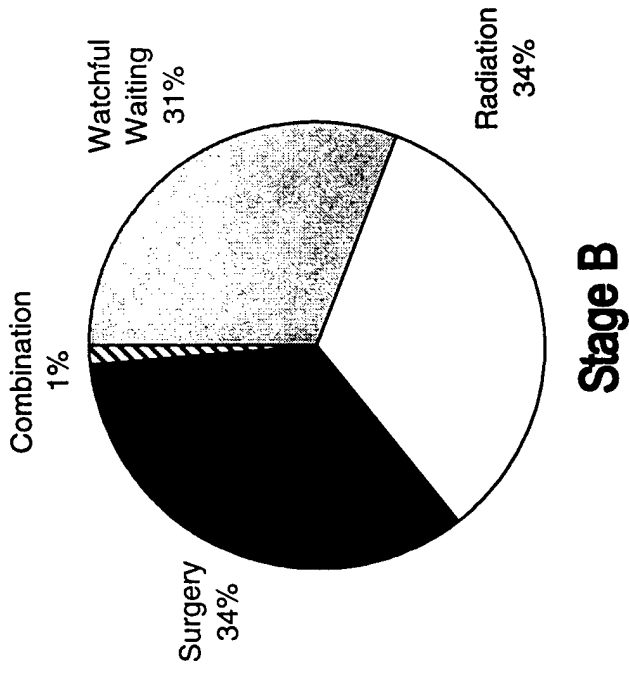
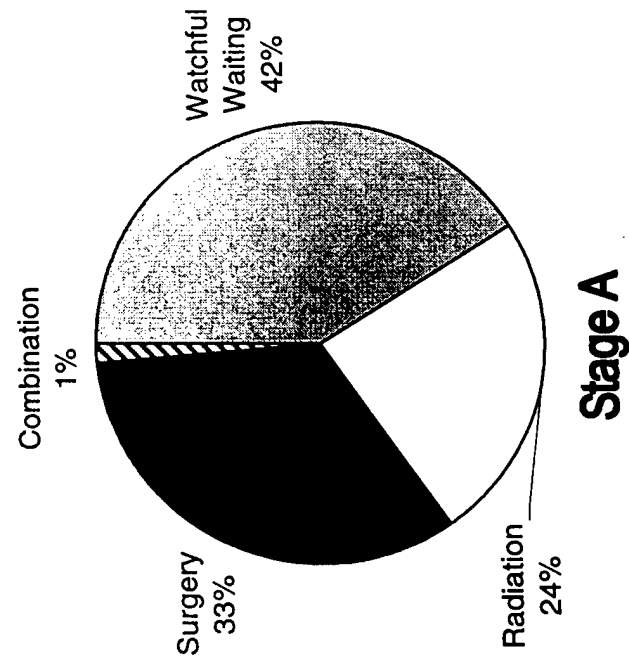


Figure 3: Prostatectomy and radiation treatment rates for prostate cancer according to age and stage defined by most accurate information from pathologic and clinical diagnoses

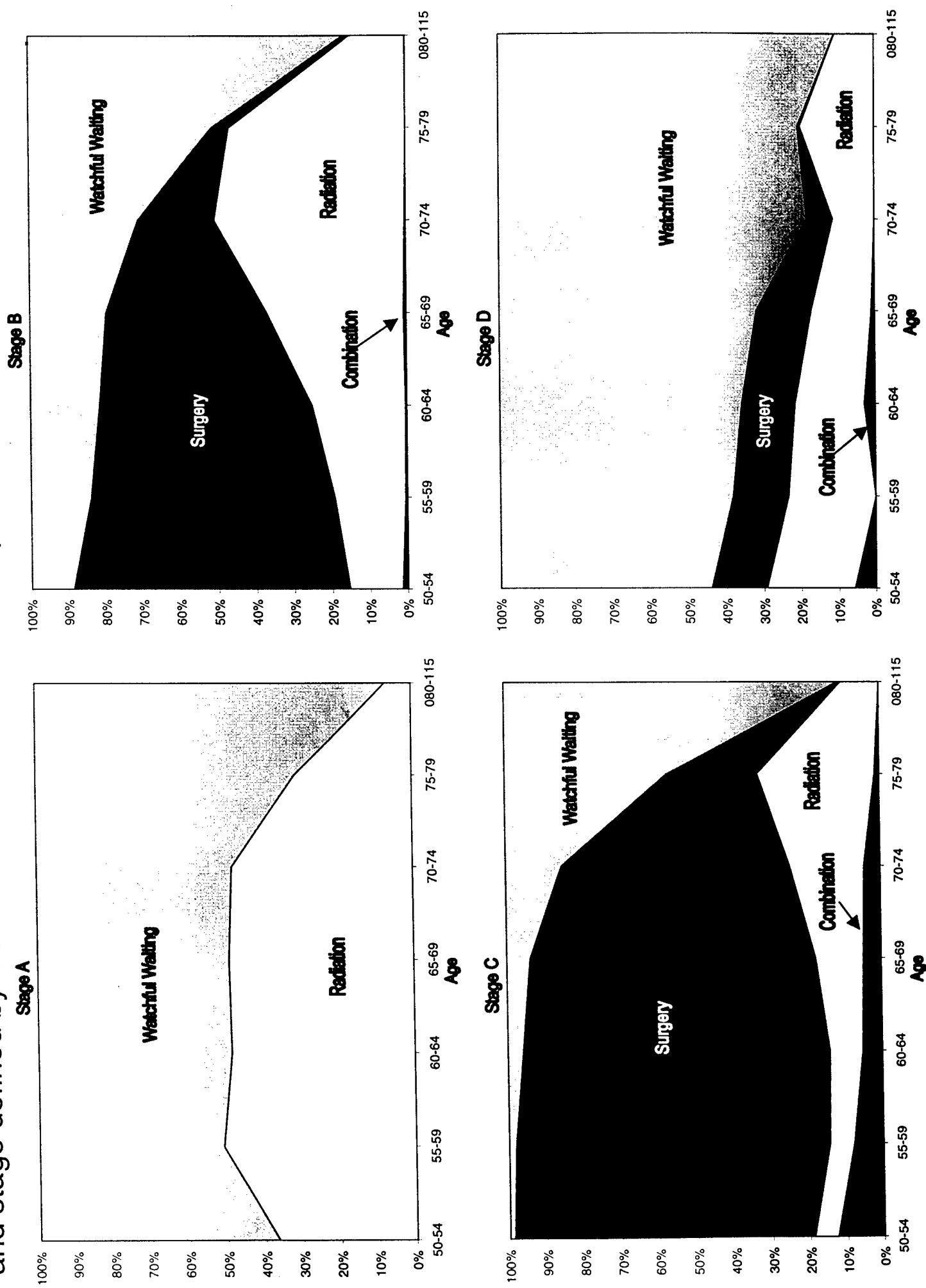
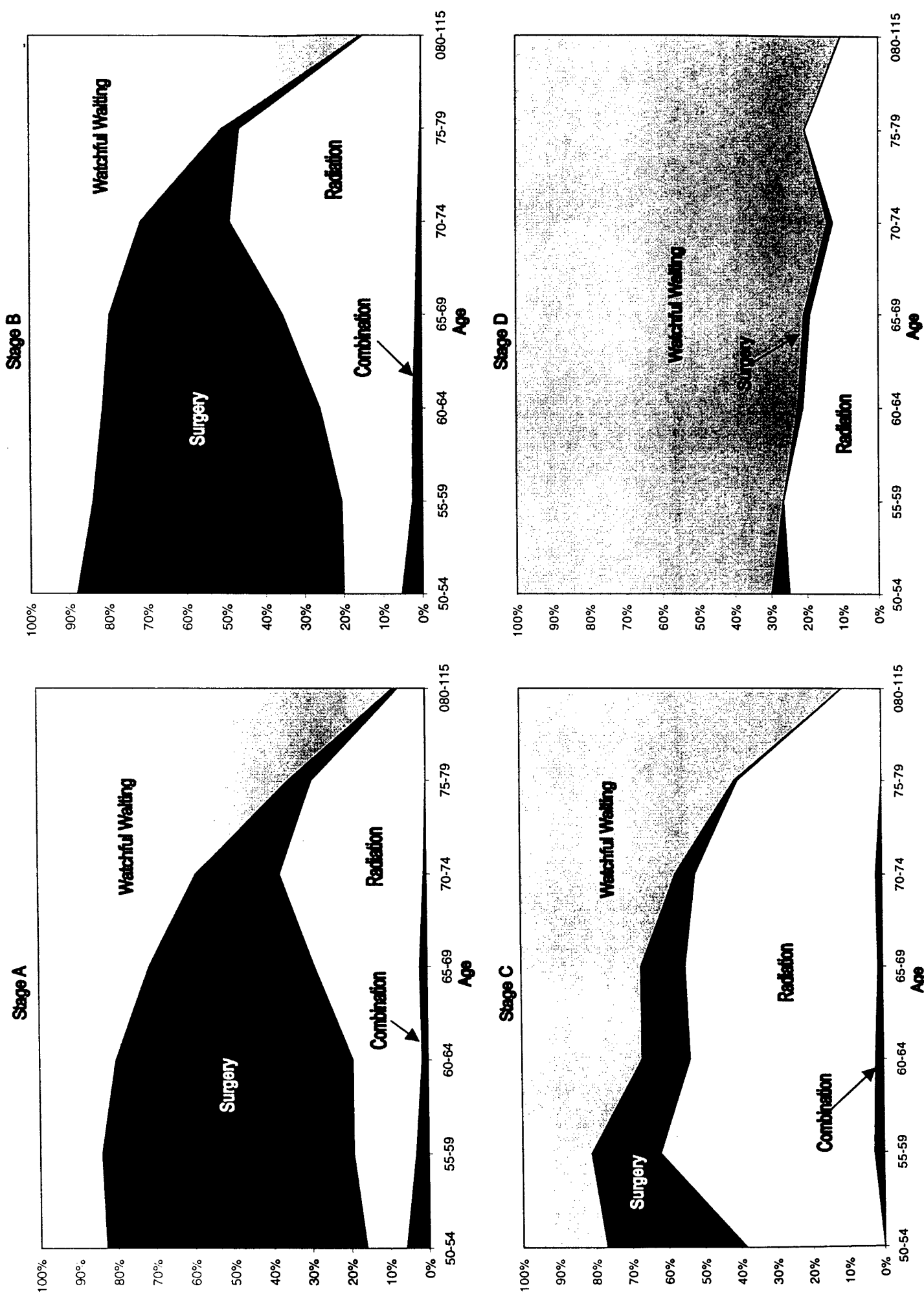


Figure 4: Prostatectomy and radiation treatment rates for prostate cancer according to age and stage defined by clinical diagnosis



Addressing Uncertainty in Medical Cost-Effectiveness Analysis:

Implications of Expected Utility Maximization for Methods to Perform Sensitivity Analysis and the Use of Cost-Effectiveness Analysis to Set Priorities for Medical Research

David Meltzer M.D., Ph.D.

Section of General Internal Medicine, Department of Economics,
and Harris Graduate School of Public Policy Studies

University of Chicago

April 6, 1999

The author gratefully acknowledges the financial support of this work by the National Institute of Aging, the Robert Wood Johnson Generalist Physician Faculty Scholars Program, and the Department of Defense Prostate Cancer Research Project. I would also like to thank Joshua Angrist, David Cutler, Sue Goldie, Zvi Griliches, Willard Manning, John Mullahy, and Milton Weinstein for helpful comments on earlier drafts of this paper.

Mailing Address: University of Chicago, 5841 S. Maryland Avenue MC 2007, Chicago, IL 60637.
Tel: (773) 702-0836 Fax: (773) 834-2238 E-mail: dmeltzer@medicine.bsd.uchicago.edu.

Abstract: Uncertainty about the benefits and costs of medical interventions make sensitivity analysis essential in medical cost-effectiveness analysis. However, there has been insufficient discussion of the objectives for performing sensitivity analysis and, therefore, the standards by which to assess methods for sensitivity analysis. This paper attempts to identify the objectives for performing sensitivity analysis in cost-effectiveness analysis and uses an expected utility model to derive methods suited to reaching those objectives. The results indicate that if the objective is to maximize expected utility given available information and if financial risk is effectively diversified through either public or private insurance, then the optimal decision is determined by the ratio of the expected cost divided by the expected benefit. If the objective of sensitivity analysis is to guide decisions for subgroups that differ from a base case, then the ratio of expected costs to expected benefits for each subgroup is the appropriate criterion. If the objective of sensitivity analysis is to set priorities for the acquisition of additional information, the incremental increase in expected utility with additional information is the appropriate measure of benefit. Though such ideal value of information calculations may be difficult to perform, certain simpler approaches to sensitivity analysis may provide bounds on the value of information with less stringent data requirements. Together, these approaches suggest methods by which the vast literature applying cost-effectiveness analysis to medical interventions can be used to help set priorities for medical research.

JEL Classification: I18; D61; O32

Keywords: Cost-Effectiveness Analysis, Uncertainty, Sensitivity Analysis, Health Care Research, Value of Information, Welfare Economics.

Section 1: Introduction

Despite some recent slowing in the growth of health care costs in the United States, health care costs have risen substantially over the past several decades and are likely to continue rising (Smith *et al.*, 1998). This appears to be largely due to the growth of new technology (Fuchs, 1990; Newhouse, 1992). While improvements in health are highly valued (Cutler and Richardson, 1997; Murphy and Topel, 1998), there is evidence from diverse methodological perspectives that many new technologies may have little value at the margin (Eddy, 1981; Brook *et al.*, 1983; McClellan, McNeil and Newhouse, 1994). Cost-effectiveness analysis and other approaches to formal medical technology assessment have arisen to attempt to address this important problem.

One of the main challenges faced by medical cost-effectiveness analysis has been the question of how to perform these analyses in the presence of uncertainty about the benefits and costs of medical interventions. This uncertainty may result either from limited evidence from clinical trials or the need to extrapolate based on the results of clinical trials using decision analyses and their associated uncertainties in the structure and parameters of decision models. This uncertainty concerning the benefits and costs of medical interventions has resulted in the devotion of a great deal of effort to performing sensitivity analyses in cost-effectiveness analyses.

However, though there have been many proposals about how to address uncertainty in cost-effectiveness analyses, there has been insufficient discussion of the objectives in performing sensitivity analysis. Without a clear understanding of these objectives, there is no compelling criterion by which to assess the merits of the many alternative approaches to sensitivity analysis.

For this reason, the lack of clarity concerning the objectives for sensitivity analysis is an important reason for the continuing ambiguity concerning methods to account for uncertainty in medical cost-effectiveness analysis.

This paper attempts to identify the objectives for sensitivity analysis within cost-effectiveness analysis and to develop methods suited to reaching those objectives. The primary objectives of sensitivity analysis are argued to be: (1) to help a decision maker to make the best decision in the presence of uncertainty, (2) to identify the sources of uncertainty to guide decisions for individuals or subgroups with characteristics that differ from a base case, and (3) to set priorities for the collection of additional information. This paper studies these problems by examining the implications of an expected utility maximization model for the optimal choice of medical interventions when there is uncertainty about the costs and benefits of those interventions. The results indicate that if the objective is to maximize expected utility given available information - as is implied, for example, by maximization of quality-adjusted life expectancy - and if financial risk is effectively diversified through either public or private insurance, then the optimal decision is determined by the ratio of the expected cost divided by the expected benefit. Other assumptions about preferences or insurance will yield other conclusions about how to account for uncertainty (Mullahy, 1997), but also would require different models for cost-effectiveness in the absence of uncertainty at the population level. These findings also have implications for sensitivity analyses done for other purposes. If the objective of sensitivity analysis is to guide decisions for subgroups that differ from the base case, then the ratio of expected costs to expected benefits for that subgroup is the appropriate criterion. If the objective of sensitivity analysis is to set priorities for the acquisition of additional information, the incremental increase in expected utility with additional information is the

appropriate measure of benefit. Though such ideal value of information calculations may be difficult to perform, certain simpler approaches to sensitivity analysis may provide bounds on the value of information with less stringent data requirements. Together, these approaches suggest a theoretically grounded approach by which the tools of medical cost-effectiveness analysis can be used to help set priorities for medical research. Following these approaches, it may be possible to draw upon the vast literature on the cost-effectiveness of specific medical interventions (Elixhauser et al., 1998) to address crucial needs for more systematic ways to set priorities for medical research. After active discussion between Congress, the Administration, and the leadership of the National Institutes of Health over the value of and priorities for Federal funding of biomedical research, the need for such systematic approaches to identify priorities for research at the National Institutes of Health was recently highlighted in a report of the Institute of Medicine (Institute of Medicine, 1998).

Section 2 discusses the objectives of sensitivity analysis. Section 3 discusses the primary methods currently used to perform sensitivity analysis. Sections 4 and 5 use an expected utility maximization model to derive methods for optimal decision making in the context of uncertainty about population outcomes. Section 6 uses this model to derive methods for sensitivity analysis to guide decisions for individuals or subgroups that differ from the base case. Section 7 applies these principles to a stylized decision concerning a medical treatment of uncertain benefit. Section 8 uses the model to derive methods for the use of sensitivity analyses to set priorities for the collection of additional information to guide decision making, including approaches to bound value of information calculations with limited information. Section 9 applies these ideas to a stylized model of the decision whether to treat prostate cancer and discusses some challenges in implementing these approaches to set priorities for research. Section 10 concludes.

Section 2: Objectives for Sensitivity Analysis

In order to begin to assess methods to account for uncertainty in population-level outcomes in cost-effectiveness analyses, it is essential to consider the objectives in performing sensitivity analyses. Although not all of these objectives may be relevant in every application, the objectives appear to fall into three broad categories: (1) to help a decision maker make the best decision in the presence of uncertainty about costs and effectiveness, (2) to identify the sources of uncertainty to guide decisions for individuals or groups with characteristics that differ from a base case, and (3) to set priorities for the collection of additional information.

(1) Decision-Making under Uncertainty about Cost and Effectiveness

This is probably the most common reason that sensitivity analysis is performed in medical cost-effectiveness analysis, and arises because the scientific literature often does not provide precise information concerning effectiveness or costs. For example, the efficacy of an immunization or the frequency and cost of complications may not be known with confidence. Nevertheless, patients may have to make decisions about whether they want the immunization and public and private insurers must decide whether they will cover it. Thus having a mechanism to help guide decision making when the costs and benefits of a medical intervention are uncertain is important.

(2) Decision-Making for Individuals or Subgroups that Differ from a Base Case

Though not frequently stated as a motivation for sensitivity analysis, developing insight into decisions faced by individuals or subgroups is also a common motivation for performing

sensitivity analysis in medical cost-effectiveness analysis. For example, a cost-effectiveness analysis for immunization of a population would likely consider the average risk of acquiring an infection in the absence of immunization. However, an analyst examining the cost-effectiveness of immunization for an individual or group with a known risk factor for acquiring some infection would want to reflect that higher-than-average risk.

(3) Priority-Setting for the Collection of Additional Information

When the conclusions of a cost-effectiveness analysis are altered by parameter values that cannot be ruled out based on the literature, the collection of additional information concerning those parameters may be justified. Though in practice it is not frequently done, sensitivity analysis can be used to identify parameters that may change the results of a decision analysis and those parameters may then be studied more intensively. A few studies have used this approach to determine the value of sample size for clinical trials (Claxton and Posnett, 1996; Hornberger, 1998), or to perform sensitivity analysis in a decision model by calculating the expected value of perfect information concerning specific parameters of the model (Felli and Hazen, 1998).

Although these three motivations for performing sensitivity analysis are clearly distinct, papers in the literature commonly do not distinguish among them in their discussion of the sensitivity analysis. This is important because different methods for sensitivity analysis may be better suited to different objectives. This is discussed further below.

Section 3: Methods for Sensitivity Analysis

Before attempting to derive methods for performing sensitivity analysis, it is useful to discuss the existing methods.

The oldest and most commonly used forms of sensitivity analyses are univariate sensitivity analyses. Following these approaches, analysts begin with the mean or modal values of all the probabilities in their analysis and use those to calculate the costs and benefits for a "base case" analysis. The parameters are then varied individually across a range of possible outcomes to see how the cost-effectiveness of an intervention changes. In some instances, the parameter values are varied over the range of all possible values, while in other cases they are varied across confidence intervals that are drawn from the medical literature.

A major advantage of one-way sensitivity analyses is that they permit the analyst to identify the effects of individual parameters on the analysis. Another advantage of one-way sensitivity analyses is that the results can be easily calculated and reported. However, there are also a number of significant shortcomings of these approaches. First, the approaches do not clearly delineate either what range of parameter values to consider or what to do when some of those possible parameter values would change the optimal decision. For example, consider again the case of a vaccination. Its probability of providing immunity is logically constrained to a number between zero and one. High and low estimates in the literature might be 0.98 and 0.60. The best study might predict a protection rate of 0.92 with a 95 percent confidence interval of 0.89 to 0.96. Which of these are we to choose in setting the range of parameters? If we choose the broadest range, it may be impossible to pin down the costs and benefits with sufficient precision to determine whether the intervention is worthwhile. If we use a 95% confidence

interval and find a benefit throughout the range, the potential for an immense harm that could occur if the true value of the parameter falls outside that range would fail to be recognized. Even if we find that the optimal decision changes for a parameter value at the upper end of the 95% confidence interval, it is not clear how that should change the decision we should make in the presence of uncertainty. If the welfare benefits over the majority of the interval are quite large, and any welfare loss at an extreme of the confidence interval is modest, it is not clear that the negative result at the extreme should have much influence on the decision made. Threshold analyses - which identify the parameter values at which an analysis crosses a cost-effectiveness threshold - are subject to the same criticism for failing to represent the magnitude of the effect of the parameter on costs and outcomes and therefore the significance of the fact that the cost-effectiveness ratio crosses some threshold for certain parameter values.

Another concern with one-way sensitivity analyses is that they may be misleading if the results obtained by varying a parameter depend on the level of other parameters in the model. This has motivated multi-way sensitivity analyses in which parameters are varied simultaneously across plausible or likely ranges. These analyses are subject to all the concerns described above concerning one way sensitivity analyses, as well as some additional problems. One problem associated with these approaches is that the number of sensitivity analyses that must be performed rises exponentially with the number of parameters. Another problem is that the probability distributions of different parameter values will generally not be independent. For example, assumptions about the natural history of untreated disease may have implications for the history of disease under treatment. This has motivated efforts to examine the joint distribution of the parameters in a model. This approach, along with a similar population-based sampling approach to estimating costs, effectiveness and cost-effectiveness ratios, sometimes

termed stochastic cost-effectiveness analysis, appears to be receiving increasing attention in the field (O'Brien *et al.*, 1994; Gold *et al.*, 1996; Polsky *et al.*, 1997). However, these analyses still do not address the question of the optimal decision in the presence of uncertainty because they do not suggest what to do when the set of possible costs and outcomes include ones that would make the cost-effectiveness ratio fail to meet the chosen threshold for cost-effectiveness.

Furthermore, there are a set of issues related to the calculation of cost-effectiveness ratios as ratios, and the relationship between those ratios and resource allocation. The ratio issue is important because benefits and costs will generally not be independently distributed since changes in parameter values will often influence both simultaneously and because ratios and their expectations are undefined when the benefits are, or may be, zero. Likewise, cost-effectiveness ratios may have very different meanings depending on the signs of cost and benefits so that merely calculating averages or even confidence intervals for cost-effectiveness ratios would not generally be meaningful (Stinnett and Paltiel, 1998). One creative approach to these issues is to reformulate cost-effectiveness analyses in terms of Net Health Benefits (Stinnett and Mullahy, 1997), in which both costs and benefits are expressed in the common denominator of years of life saved. While free of some of the complications associated with estimating cost-effectiveness ratios, the usefulness of the Net Health Benefit approach is diminished by the fact it does not allow easy comparisons with results from traditional cost-effectiveness analyses that rely on cost-effectiveness ratios, and is dependent on assumptions about the valuation of improvements in health. A related approach with similar concerns is to convert health benefits into a monetary value, as is done in cost-benefit analysis (Tambour, Zethraeus, and Johannesson, 1998).

In assessing the methods described in this section, it is interesting to note that while all of them appear to have some significance for the objectives described in the preceding section, none of them are explicitly linked to that set of objectives. As discussed above, this lack of clarity concerning the objectives for sensitivity analysis is an important reason for the continuing ambiguity concerning methods to account for uncertainty in medical cost-effectiveness analysis. The next two sections use an expected utility maximization model to attempt to develop an approach to assess the importance of uncertainty about parameter values in order to make an optimal decision under uncertainty. The sections that follow then examine the adaptation of that approach to address the other two common objectives of sensitivity analysis – the determination of cost-effectiveness for individuals or subgroups and the identification of areas where the collection of additional information would be of value.

Section 4: A Deterministic Model of Health Outcomes with Uncertainty about Effectiveness

In this simple case, we assume that there is uncertainty about the effectiveness ($\theta \in \Theta$) of a medical intervention (m) but that the outcome of a medical intervention given θ is certain. By making this assumption, we abstract from the problem of uncertainty in outcome for an individual, and focus instead on uncertainty for a “representative consumer”. To capture the possibility that effectiveness may affect both the costs and benefits of an intervention, we allow both utility (U) and the costs of the medical intervention (c) to depend directly on θ . In addition, utility is assumed to depend on non-medical consumption (x) and medical expenditure. Thus, $U = U(m, \theta, x(\theta))$ and $c = c(m, \theta)$. Here x is written as $x(\theta)$ to denote the fact that x will vary with

θ for any m to satisfy the budget constraint $c(m, \theta) + x(\theta) - I = 0$ for each level of effectiveness. To model cost-effectiveness, we assume that people maximize expected utility¹ and take the example of a representative consumer who maximizes expected utility subject to a budget constraint conditional on each level of effectiveness:

$$\underset{m}{Max} \int p(\theta) U(m, \theta, x(\theta)) d\theta \text{ such that } c(m, \theta) + x(\theta) - I = 0 \text{ for all } \theta \quad (1)$$

This can then be rewritten as a LaGrange multiplier problem in which the $\lambda(\theta)$ is the multiplier for the budget constraint for each level of θ , which are multiplied by $p(\theta)$ without loss of generality to yield:

$$\underset{m, \lambda(\theta)}{Max} \int p(\theta) U(m, \theta, x(\theta)) d\theta + \int \lambda(\theta) p(\theta) [I - c(m, \theta) - x(\theta)] d\theta \quad (2)$$

This generates a first order condition for medical expenditure which is:

$$\int p(\theta) \frac{\partial U(m, \theta, x(\theta))}{\partial m} d\theta + \int \lambda(\theta) p(\theta) \frac{\partial c(m, \theta)}{\partial m} d\theta = 0 \quad (3)$$

This condition implies that investment in a medical intervention should occur to the point at which its expected marginal benefit (utility) equals the expected value of the marginal-utility-of-income-weighted marginal cost. Permitting the marginal utility of income to be dependent on θ reflects the possibility that, either because of changes in the utility function or changes in costs due to θ , income might have a greater or lesser marginal utility.

For an individual, these effects of uncertainty about the costs and effectiveness of medical interventions on the marginal utility of income are clearly plausible and potentially important. If someone has hip replacement for arthritis at age 55 and then suffers a severe complication, is forced into early retirement, and requires around-the-clock care, both their

¹ While individual preferences may in fact be inconsistent with expected utility maximization, QALYs implicitly assume that people maximize expected utility. While the relaxation of this assumption might be desirable, doing so

utility and medical costs will be directly affected and their marginal utility of income could change substantially. In a population, however, such effects are far less compelling because insurance can equate the marginal utility of income across health states unless an intervention leads to an extraordinarily large change in either population health or costs. Thinking from a population perspective in which most extremely expensive medical interventions affect a relatively small number of persons and most common medical interventions are relatively modest in cost, it is much less likely that the (aggregate) marginal utility of income will change substantially with uncertainty about the costs or benefits of a single intervention². If this is the case, then $\lim \lambda(\theta) \rightarrow \lambda$ and the first order condition for medical expenditures converges to:

$$\int p(\theta) \frac{\partial U(m, \theta, x(\theta))}{\partial m} d\theta + \int \lambda p(\theta) \frac{\partial c(m, \theta)}{\partial m} d\theta = 0 \quad (4)$$

which implies that the cost-effectiveness ratio is:

$$\frac{\int p(\theta) \frac{\partial c(m, \theta)}{\partial m} d\theta}{\int p(\theta) \frac{\partial U(m, x(\theta))}{\partial m} d\theta} = \frac{1}{\lambda} \quad (5)$$

Thus expected utility maximization implies that the optimum cost-effectiveness ratio of an intervention in a population under uncertainty is closely approximated by the ratio of expected costs to expected benefits. Note that this “ratio of means” solution is analogous to that suggested by Stinnett and Paltiel (1997) as the solution to a constrained optimization problem in a linear programming context and by Claxton (1998) in a Bayesian discrete choice decision theoretic

would therefore involve a substantial reformulation of the way in which health benefits are assessed even in the absence of uncertainty. This is discussed further in the conclusion of this paper.

² Note that even if changes in health status were to lead to substantial changes in income or the need for non-medical assistance holding income constant across individuals in different health states, optimal insurance could still equate the marginal utility of income across states. In practice, of course, insurance will generally fall short of this ideal, but this is nevertheless a useful point of reference.

context. However, neither analysis derives the result directly from a formal utility maximization model nor addresses the possible dependency of the marginal utility of income on θ .

While this argument about the dependence of the marginal utility of income on θ has not been made previously within the context of medical cost-effectiveness analysis, it should be noted that the argument is quite similar to that made by Arrow and Lind (1970) concerning the evaluation of risk in public investment decisions. In that article, the authors argue that the large scale on which the public sector exists allows it to effectively eliminate any welfare loss associated with the riskiness of investments by spreading the risk across a sufficiently large population. The argument here relies both on this diversification effect and the relatively modest magnitude of almost any one public health care decision in the context of overall health and health expenditures.

Section 5: A Stochastic Population Model with Individual-level Uncertainty about Outcomes

Unlike the deterministic model presented above, actual medical interventions almost always have uncertain outcomes for individuals even for any given set of population parameters. Thus if there is a set of individuals indexed by $j \in J$ who might each experience health outcome $\varepsilon_j \in E$, the probability of experiencing outcome ε_j given $\theta \in \Theta$ can be written as $f(\varepsilon_j | \theta)$ and expected utility can be written as:

$$\int p(\theta) \int \int f(\varepsilon_j | \theta) U_j(m, \varepsilon_j, x_j(\varepsilon_j, \theta)) d\varepsilon_j dj d\theta \text{ such that}$$

$$c_j(m, \varepsilon_j, \theta) + x_j(\varepsilon_j, \theta) - I = 0 \text{ for all } \theta, j, \varepsilon_j \quad (6)$$

Following the lines of the argument above, we can construct state-specific LaGrange multipliers $\lambda_j(\epsilon_j, \theta)$ and note that if there is (1) a large population so that aggregate risk given θ is negligible, (2) full insurance, and (3) uncertainty in the effectiveness of the intervention has limited consequences in the sense that θ does not have much effect on λ as described above, then

$$\lim \lambda_j(\bar{\epsilon}, \theta) \rightarrow \lambda(\bar{\epsilon}, \theta) \rightarrow \lambda \text{ for all } \bar{\epsilon} \equiv \{\epsilon_1, \dots, \epsilon_j, \dots, \epsilon_n\} \text{ where } \epsilon_j \in E \text{ and for all } \theta. \quad (7)$$

This then implies again that cost-effectiveness can be identified by the ratio of expected costs to benefits.

Section 6: Sensitivity Analysis to Guide Individual or Subgroup Decisions

When sensitivity analysis is done to guide decisions for individuals or subgroups, the problem is essentially the same as for the total population, except that some of the parameters in the parameter vector θ have a different probability distribution $p'(\theta)$ than in the overall population. This may arise if parameters for those individuals or subgroups are known to differ from those for the population as a whole. It may also arise if prior probabilities concerning the distribution of parameters for sub-populations differ from what they are in the population as a whole. This could be the case either if the true parameters differ in the populations or if they are the same but some populations are more or less well studied than others. In either case, the analysis differs only in the probability distribution for the parameters. Accordingly, the solution to this problem for individuals or subgroups is again the ratio of the expected value of costs to the expected value of benefits, only using the appropriate prior probability distribution for the subgroup or individual.

Section 7: Application to a Stylized Decision Concerning a Treatment of Uncertain Benefit

Figure 1 describes a stylized decision concerning an intervention of uncertain benefit. For simplicity, the intervention is assumed to cost \$10,000 with certainty. Uncertainty is assumed to exist only with respect to benefits; it is assumed that there is a 90% chance that the benefit is 0.1 life-year, but also a 5% chance each that the benefit is 0.01 or 1 life year.

Taking these three possibilities individually, the cost-effectiveness ratios are \$100,000, \$1,000,000, or \$10,000 respectively. If one used a cutoff of \$100,000 per life-year, a traditional sensitivity analysis would therefore be indeterminate. Such indeterminacy is, in fact, an extremely common result in cost-effectiveness analyses. Another limitation of this standard approach is that, while the cost-effectiveness ratios tell us something about the magnitude of the benefits relative to the cost, they do not provide any indication of how to relate the magnitude of those benefits to their likelihood.

Common approaches to sensitivity analysis might take other perspectives. For example, the stochastic cost-effectiveness approach might conclude that since there is only a 5% chance that the intervention is not cost-effective, it should be selected. On the other hand, the same approach could be used to argue that since there is only a 5% chance that the intervention will provide a benefit in excess of its cost, it should not be selected. The problem with these perspectives is that they do not reflect the magnitude of potential benefits relative to costs.

Following the expected utility approach described above, the expected cost is \$10,000 and the expected benefit is: $0.05 \times 0.01 + 0.9 \times 0.1 + 0.05 \times 1.0 = 0.0005 + 0.09 + 0.05 = 0.1405$ life-years. Thus the cost-effectiveness ratio is $\$10,000 / 0.1405 \text{ life-years} = \$71,174/\text{life-year saved}$, which is clearly cost-effective by the \$100,000 per life-year standard. Even though the chance

that the intervention is highly beneficial is only 5%, more than one-third ($0.05/0.1405=36\%$) of the expected benefit comes from the unlikely event that it is highly effective. It is this ability to incorporate both the magnitude and likelihood of benefits and costs into a single statistic that can be used to guide decision-making that is the primary advantage of the expected value approach over the traditional approaches that incorporate only one or the other dimension, and often result in indeterminate conclusions that do not provide much guidance for decision making.

Section 8: Sensitivity Analysis to Guide Information Collection

In addition to providing guidance about how to identify the optimal decision under uncertainty given available information, the expected utility approach can be used to set priorities for research by assessing whether the collection of additional information is likely to be worthwhile. When a study is done to accumulate improved information concerning parameters in a decision model, the value of information is the change in expected utility that comes from a change in uncertainty about the parameters. Although this fundamental principle dates back at least to the pioneers of statistical decision theory (e.g. Raiffa and Schlaifer, 1961; Pratt, Raiffa, and Schlaifer, 1965), it has not been commonly used in developing techniques for sensitivity analysis in medical decision analysis. Indeed, when formal techniques for clinical trial design have been implemented (e.g. O'Brien *et al.*, 1994; Al, van Hout *et al.*, 1998; Briggs and Gray, 1998), they have often been based on criteria for decision making such as confidence intervals around the cost-effectiveness ratio, which generate suboptimal results for the same reasons as the problems with threshold approaches to sensitivity analysis that are described above. Two exceptions to this are Claxton and Posnett (1996) and Hornberger (1998), which focus on the

determination of optimal sample size for a clinical trial from a cost-effectiveness perspective in a full bayesian context.

Adopting the expected utility approach, assume that for any information set describing the parameter distribution, $p(\theta)$, there is an optimal choice of m as described above. Call this $m^*(p(\theta))$. This implies an expected utility with existing information (EU_0) of:

$$\int p(\theta) U(m^*(p(\theta)), \theta, x(\theta)) d\theta \quad (8)$$

Now imagine that we are able to acquire additional information about θ . Assume further that the cost of this research is c_r . Though the analysis is easily generalized to permit an infinite number of possible outcomes of the experiment³, assume for simplicity that there are only two possible outcomes of this experiment: with probability q that the distribution of θ is found to be $p'(\theta)$ and with probability $(1-q)$ that it is found to be $p''(\theta)$, where, for consistency with the initial prior distribution, $q \cdot p'(\theta) + (1-q) \cdot p''(\theta) = p(\theta)$. In these cases, the optimal level of medical expenditure will be $m^*(p'(\theta))$ and $m^*(p''(\theta))$ and the expected level of utility is:

$$q \int p'(\theta) U(m^*(p'(\theta)), \theta, x^{*'}(\theta)) d\theta + (1-q) \int p''(\theta) U(m^*(p''(\theta)), \theta, x^{*''}(\theta)) d\theta \quad (9)$$

where $x^{*'}(\theta)$ and $x^{*''}(\theta)$ are determined from the budget constraint net of research costs c_r (i.e.

³ In the general case, we wish to compare the expected utility resulting from the optimal decision m^* given the original budget constraint in the absence of information to the expected utility resulting from the optimal decision in the presence of the new information subject to a budget constraint that includes the cost of collecting information (C_r). Thus we compare the expected utility resulting from the solution to:

$$\text{Max}_{m, \lambda(\theta)} \int p(\theta) U(m, \theta, x(\theta)) d\theta + \int \lambda(\theta) p(\theta) [c(m, \theta) + x(\theta) - I] d\theta$$

to the expected utility of the solutions to the optimal decision problems for the j possible outcomes of the experiment as individually given by

$$\text{Max}_{m, \lambda(\theta)} \int p_j(\theta) U(m, \theta, x(\theta)) d\theta + \int \lambda(\theta) p_j(\theta) [c(m, \theta) + x(\theta) + c_r - I] d\theta \quad \text{where} \quad \int_{j \in J} p_j(\theta) p(j) dj = p(\theta).$$

$c(m, \theta) + x(\theta) + c_r - I = 0$ for all θ . It follows that the change in expected utility with the collection of information, or expected value of information (EVI) is:

$$q \int p'(\theta) U(m^*(p'(\theta)), \theta, x(\theta)) d\theta + (1-q) \int p''(\theta) U(m^*(p''(\theta)), \theta, x(\theta)) d\theta - EU_0 \quad (10)$$

If this is positive then the study is worth performing, if not, then it should not be performed.

Although this value of information calculation is easily described in theoretical terms, the ability to implement this approach depends on the ability to assemble meaningful information on the prior probabilities of the parameters required for the calculation, and these may be very difficult to obtain. In some instances, these may be estimated based on published studies that report means and confidence intervals that may then be used to describe the full distribution of parameter values. In other instances, primary data collection may be required. Still, it is likely that in a significant number of cases it will not be possible to identify much information that will inform priors. Moreover, it may be quite difficult to say much about how an experiment is likely to affect the posterior distributions of the parameters.

These empirical challenges suggest that techniques for assessing the value of information that do not rely on this data concerning prior or posterior distributions would be highly useful. In the case where information on priors is available, one such possibility is the expected value of perfect information: $EVPI = \int p(\theta) U(m^*(\theta)) d\theta - EU_0$, where $m^*(\theta)$ is the optimal choice of m if θ is known. Since the expected value of information is always positive⁴, this provides an

⁴ To see this, note that if research cost are zero, the fact that $m^*(p'(\theta))$ and $m^*(p''(\theta))$ are optima implies that the first two terms in this equation are greater than:

$$\begin{aligned} & q \int p'(\theta) U(m^*(p'(\theta)), \theta, x(\theta)) d\theta + (1-q) \int p''(\theta) U(m^*(p''(\theta)), \theta, x(\theta)) d\theta \\ & = \int p(\theta) U(m^*(p(\theta)), \theta, x(\theta)) d\theta, \end{aligned} \quad (11)$$

which is the expected utility from the optimal expenditure level in the absence of information. This implies that the expected value of free information is greater than zero. For completeness, it should be noted that this result applies

upper bound on the ideal value of information calculations described above. From a practical point, however, the advantage of the EVPI calculation is that it does not depend on the posteriors. Indeed, this is probably one reason why the EVPI approach has been used in the cost-effectiveness literature (e.g. Felli and Hazen, 1998)⁵.

Although EVPI is simpler to determine than EVI, it still depends on knowledge of the priors. An alternative measure that did not depend on this might also be useful. One such measure is the maximal value of information (MVI) over all possible values of $\theta \in \Theta$, $MVB = \max_{\theta \in \Theta} U(m^*(\theta))$. Although this will also only be an upper bound on EVPI and, therefore, EVI, it depends only on knowledge of the value function conditional on θ . Despite the fact that it may be a relatively crude upper bound, it is worthwhile to note that this criterion in fact corresponds to that implied by a threshold analysis in which the bounds are determined by the extreme values of the parameter (assuming, as is usually done, that the value function is monotonic with respect to the parameters). Thus, application of the threshold technique based on the full range of possible values of a parameter can be considered a bound on the more general value of information calculation, only with less rigorous information requirements. Thus, like

only to public information. When information is private, there is no guarantee that its expected value will be positive (e.g. Rothschild and Stiglitz (1976)).

⁵ It should be noted, however, that Felli and Hazen (1998) consider the expected value of perfect information relative to the expected value of an optimal decision that they specify as one that maximizes the expected payoff given parameter values that the investigator feels are "most likely to obtain" (p. 100). This seems to suggest the modal value(s) of the parameter(s). Nevertheless, in their applications they tend to choose the mean values of their parameters. Regardless, since neither of these are generally the optimal decision given available information, this calculation will overstate the actual expected value of perfect information relative to an optimal decision given imperfect information. In this sense, the calculation by Hazen and Felli can be viewed as an upper bound on the true EVPI. The only advantage of this approach over the theoretically correct approach is that it avoids the need to determine the decision that maximizes the expected value with existing information. This is not a substantial advantage, however, because, if the value of information is small, then collecting additional information is presumably not worthwhile and therefore knowing the optimal decision with existing information is key. Similarly, if the value of information is large, then one still wants to try to determine the EVPI relative to an optimal decision with existing information in order to see how much that decreases the bound on the EVPI. Thus, in either case, the calculation of EVPI relative to the optimal decision given current information is preferred. It is also generally not an

EVPI, the threshold technique based on the full range of possible values a parameter might take can be considered a method for placing an upper bound on the more complex EVI calculation. When these calculations suggest the MVI or EVPI is low, the full EVI calculation is not necessary. Note, in contrast, that the common practices of assessing cost-effectiveness at a 95% confidence interval for a parameter or calculating stochastic cost-effectiveness intervals have no clear theoretical justification.

Thinking more broadly, if Θ is enlarged to include any possible conceivable value of θ , even if those values are excluded under current technology, this type of reasoning can be extended to consider any possible research on the parameter in question. For example, if the probability of cure with the best current treatment for a disease is known to be between 20 and 40 percent with certainty and the treatment is found not to be worthwhile (perhaps because of morbidity), one could calculate whether treatment would be worthwhile if the cure rate were 100 percent. This might be called the maximum value of research (MVR), and, in turn, can be used to generate an upper bound on MVI that does not require any data at all concerning the parameter in question. The MVR concept could also be expanded to consider innovations that led to fundamental changes in the structure of the decision tree, and not just the effects of changes in its parameters.

Section 9: Application to a Stylized Model of the Decision whether to Treat Prostate Cancer

In order to illustrate the approaches described above, this section examines a simplified model of the decision to treat prostate cancer. A highly stylized model is chosen in order to

extremely difficult determination to make once the ability to assess the expected value of outcomes from an arbitrary (optimal or suboptimal) decision is present.

focus attention on the methods described rather than the specific application. In this simplified model (Figure 2), the decision to treat prostate cancer is viewed as a decision between radical prostatectomy (surgical removal of the prostate) and “watchful waiting” (no intervention unless the disease is found to spread). This decision is represented by the two decision nodes in the middle of the decision tree in Figure 2. In this simplified model, radical prostatectomy is assumed to be curative, so that the patient lives out a normal life expectancy of 25 years. Radical prostatectomy is major surgery, however, and carries the risk of immediate death, which is assumed to occur 5% of the time. The outcome of watchful waiting will depend on how quickly the prostate cancer progresses. Many cancers will progress slowly enough that men die of other causes before they die of prostate cancer and thus live a “normal” life expectancy (assumed here to be 25 years). Other men will progress rapidly and die of prostate cancer (assumed here to occur at 10 years). For simplicity, we assume that quality of life is not a concern in prostate cancer treatment so that outcomes are measured in life years, which are the same as quality-adjusted life years. Radical prostatectomy is assumed to cost \$10,000 and the basic future costs of survival are assumed to be \$20,000 per year. (See Meltzer (1997) for a justification for including costs of this nature).

However, the natural history of prostate cancer is not as well understood as suggested by these assumptions. In fact there is a great deal of uncertainty even about average rates of progression to death from prostate cancer – i.e. how aggressive the disease is on average. This is the dimension of uncertainty on which we focus in this example. This is captured in a stylized way in Figure 2 by the upper and lower decision trees that differ in the fraction of tumors that will progress rapidly (0.085 in the “non-aggressive” case, and 0.2 in the “aggressive” case).

Though of course some other fraction in between or near these two might be imagined to be correct, we assume for simplicity here that one or the other of these values is precisely correct.

Panels 1 and 2 of Table 1 show the results of a cost-effectiveness analysis of the treatment decision in the non-aggressive and aggressive cases. In both cases, treatment provides a benefit, but in the first case it is a small benefit with a cost per QALY of \$420,000 and in the second case it is a much larger benefit with a cost per QALY of only \$26,000. If we assume for simplicity that the cutoff for cost-effectiveness is \$100,000 per QALY, then the optimal decision in the first case would be watchful waiting, while in the second case it would be treatment.

The left most part of the decision tree reflects the fact that we do not know which of these two possibilities might be the case and places some prior probabilities on the two arms (0.2 aggressive, 0.8 non-aggressive). Panel 3 of Table 1 reports the expected benefits and costs of the screening decision when these priors are held. In that case, the ratio of the expected costs to expected benefits is \$47,000, which would be cost-effective by the \$100,000/QALY standard. This might be considered surprising because the assumption was that there was an 80 percent chance that the progression was not aggressive, and treatment is not even close to being cost-effective by the \$100,000/QALY standard in that case. The result is driven by the 20 percent chance that the benefit could be much larger, despite the fact that that possibility is not very likely. This points out the potential for the ratio of the expected value approach to generate different results than approaches that do not account for both the magnitude and likelihood of the potential benefits.

We now turn to the question of whether the collection of additional information would be of value. Following the approach described above, we begin with calculation of the maximum value of information. This calculation can be done in a variety of ways requiring progressively

more information. To take an extreme example, assume that we knew nothing about the probability that prostate cancer is aggressive, but only the life expectancy of patients with aggressive cancers who are treated or not treated, and the price of prostatectomy. In the absence of knowledge about the probability that cancers would progress rapidly, there is no clear guidance about whether watchful waiting or prostatectomy dominates, so we therefore consider both cases as reference cases. Assume first that we decide that no treatment will be the reference point. To get an upper bound on the value of information, one could use only information on the life expectancy of treated and untreated patients and assume that all patients have aggressive cancers. Specifically, assuming that all men who have prostate cancer but are not treated live 10 years (QALYs), while those who are treated live 25 years (QALYs), the value of treatment per patient would be $15 \text{ QALYs} \times \$100,000/\text{QALY} = \$1.5 \text{ Million per patient}$. Alternatively, we could assume that that treatment would be the reference case, in which case the benefit of finding out that treatment was not cost-effective would be the cost savings from avoiding prostatectomy ($\$10,000/\text{patient}$) and avoidance of treatment-related mortality ($0.05 \text{ mortality} \times 25 \text{ years maximum life expectancy} \times \$100,000/\text{QALY} = \$125,000/\text{patient}$), which add to $\$135,000$ per patient.

If one were then to use this knowledge of the maximum value of information for a patient to make a decision about whether investment in a study to resolve the ambiguity about the aggressiveness of prostate cancer would be worthwhile, one might multiply these numbers by the number of men who are found to have prostate cancer annually (100,000) and divide by some real interest rate (0.03) to reflect the discounted value of the value of that information over time to get the maximum value of information (MVI): $\$1.5 \text{ Million} \times 100,000 / 0.03 = \5 Trillion if the baseline strategy is watchful waiting and $\$0.135 \text{ Million} \times 100,000 / 0.03 = \450 Billion if the

baseline strategy is prostatectomy. These extremely large estimates of the maximum value of information suggest that there is the potential for information of immense value to come from knowledge about the efficacy of prostate cancer treatment. This value of information is large relative to the cost of any conceivable clinical trial.

Of course these maximum value of information calculations represent an upper bound, and a fair interpretation of these findings is that the MVI approach is simply not informative in this case, despite its analytical simplicity and independence of assumptions about the probabilities that cancers are aggressive. This suggests that it is worthwhile to pursue the expected value of perfect information (EVPI) approach.

The EVPI approach is described in panel 4 of the table. The panel describes the expected value of three strategies: watchful waiting, radical prostatectomy, and the optimal decision with perfect knowledge of the actual average progression rate (EVPI). The last two columns calculate the value of the change in QALYs (assuming \$100,000 per QALY for illustration) and the net incremental benefit of the policy choice compared to the strategy immediately above it in the table.

The first point to note is that if one made policy based on the most likely cost effectiveness ratio (\$420,000), one would choose watchful waiting, but if one chose based on the ratio of the expected values, one would choose radical prostatectomy, which yields a net benefit of $(\$26,000 - \$6,400 =) \$19,600$ per patient relative to watchful waiting. This is a quantified measure of the expected gain from the improvement in decision making by using the mean of the expected values as opposed to basing the decision on the most likely cost-effectiveness ratio, as is generally done in the "base case" reported by most current cost-effectiveness analyses.

The second point to note is that the expected value of the gain versus watchful waiting with improved information is even higher at \$26,000 per patient. This implies an additional gain of \$6,400 per patient of the improved information compared to the best possible decision with the initial information. Converting this patient level estimate of the value of research into a population level estimate as above suggests an EVPI of $\$6,400 \times 100,000 / 0.03 = \21 billion. As with the MVPI, this large EVPI suggests that the expected value of perfect information about the efficacy of prostate cancer treatment would indeed be quite large relative to the cost of almost any conceivable clinical trial.

Of course this too is an upper bound on the expected value of information that would come from any actual clinical trial, since any clinical trial is likely to provide less than perfect information. Panel 5 examines one such case in which an experiment is done that has two possible outcomes: a 50 percent probability of an outcome that suggests that the probability that prostate cancer is aggressive is 0.05 and a 50 percent probability of an outcome that suggests that prostate cancer is aggressive is 0.35. (Note this preserves the prior that the probability that prostate cancer is aggressive is 0.2 since $0.5 \times 0.05 + 0.5 \times 0.35 = 0.2$.) The expected value of outcomes from watchful waiting and radical prostatectomy given these two possible outcomes of the experiment are reported in the upper and lower parts of panel 5. In the first case, the optimal decision switches to watchful waiting as compared to prostatectomy with the initial information, which yields a net surplus of \$600 per patient. In the second case, prostatectomy remains the optimal choice, so there is no additional benefit to having done the study. Thus the expected net benefit is $0.5 \times \$600 = \300 per patient. A decision about the study might be made by comparing its cost to the expected value of the information (EVI): $\$300 \times 100,000 / 0.03 = \1 billion.

Therefore, the value of this study would be quite large, although substantially less than the upper bound suggested by the EVPI.

In a similar manner, possible experiments concerning all other dimensions of the model might be examined to determine whether they would be worthwhile. In this way, it might be determined how much could be gained by improved sensitivity and specificity of screening tests, decreased complications of treatment, improved risk stratification prior to treatment, and so on.

Clearly, there has been no attempt in this example to suggest that a comprehensive attempt to perform a precise calculation of the type described would generate results anything resembling these in magnitude. However, these simplified calculations do illustrate the types of calculations that might be performed to assess the value of research, including more simple calculations such as the EVPI that require less information.

Section 10: Conclusion

This paper has examined the purposes for which sensitivity analysis is performed in medical cost-effectiveness analysis and the implications of an expected utility maximization model for the methods to perform such analyses. When sensitivity analysis is done to help guide decision making in the presence of uncertainty, expected utility maximization implies that the optimal decision is described by the ratio of the mean cost divided by the mean benefit. When sensitivity analysis is done to make decisions for subgroups with identifiable differences in costs or benefits, the standard is the same, only using the appropriate mean costs and mean benefits for each group. When sensitivity analysis is done to set priorities for the collection of additional information, the expected utility maximization framework suggests that value of information calculations can be done in the same context to assess the expected value of information (EVI) obtained by research. Though the data requirements for such calculations are substantial, the expected value of perfect information (EVPI) and maximal value of information (MVI) can be used to bound EVI with less stringent data requirements.

At a theoretical level, there are a number of limitations of this work. First, even with complete certainty about costs and benefits, cost-effectiveness analysis may not maximize the welfare of individuals or society. Reasons for this include differences between individual and social perspectives that complicate the measurement of costs and benefits (Meltzer and Johannesson, 1998), failure to measure benefits in a way that reflects the preferences of patients (Meltzer *et al.*, 1998), and fundamental issues concerning the potential to derive any consistent social preference ordering that is based on the preferences of individuals (Arrow, 1951). Also, there are important issues about how risk at the individual level may affect welfare (Kahneman

and Tversky, 1979) that are essentially ignored by the assumptions that perfect insurance is available and that people maximize expected utility. Even though the expected utility maximization assumption is implicit in QALYs, it may not well reflect the way in which people incorporate uncertainty into their preferences (Kahneman and Tversky (1979)). However, while this is clearly an important limitation of QALYs, it is one that needs to be addressed regardless of the issues about aggregate uncertainty that are addressed in sensitivity analysis. As discussed above, concerns about aggregate financial and health risk do seem somewhat less compelling in a social context where the risks associated with decisions about individual technologies are likely to be modest from an aggregate perspective. Nevertheless, the issue of how risk should be assessed in policy decisions is important and deserves further consideration because other assumptions about preferences concerning risk or about insurance would lead to different conclusions about a range of methodological issues in cost-effectiveness analysis, including sensitivity analysis (e.g. Mullahy, 1997).

Rather than using either expected utility or some other framework to incorporate preferences concerning outcomes and risk, it might be argued that it would be preferable to simply describe the joint distribution of benefits and costs. Nothing about this analysis suggests that should not be presented. Indeed, it is the primary data with which expected value calculations are performed. However, using such information to make choices would still require decisions about how to incorporate risk into decision making. Unlike traditional approaches to sensitivity analysis, the expected value approach provides direct guidance in the common case in which the optimal decision varies with the assumptions that are made.

At an empirical level, there are important challenges in developing meaningful priors concerning the parameters of decision models (e.g., probabilities, quality of life values, discount

rates, etc.). As discussed above, this may often require extensive review of existing data, primary data collection, or even sometimes analyses based on a variety of arbitrary priors. It may also be very difficult to determine priors for the likelihood that a research project will find a meaningful result. Whether it is possible to adequately address these challenges will be resolved only through efforts to apply these ideas empirically.

In thinking about using these approaches to assess the value of research, there are also a variety of additional challenges. These include the interdependence of the benefits of research projects on related topics, the possibility that the research might become less (or more) valuable over time if other technological or demographic changes arise that alter the management, frequency or natural history of the disease, and the unpredictability of how the results of research (particularly basic research) might be useful in areas outside the initial areas of inquiry. The difficulty of these issues implies that the sort of formal analyses suggested here are more likely to be of use for evaluating clinical research than for evaluating basic research.

Despite the theoretical and empirical issues raised by this work, the importance of making good decisions about the allocation of resources to medical interventions and medical research suggest that work in this area be an important priority. In considering whether formal approaches based on the tools of cost-effectiveness analysis are likely to have much influence on priority setting, it is encouraging that the recent report of the Institute of Medicine on improving priority setting at the NIH recommended that: "In setting priorities, NIH should strengthen its analysis and use of health data, such as burdens and costs of diseases, and on data on the impact of research on the health of the public. (p. 11)" (Institute of Medicine, 1998).

On the other hand, the limited number of cases where cost-effectiveness analysis has been highly influential in medical resource allocation decisions and the likely resistance of the

medical research community to having research proposals evaluated by formal criteria suggest that formal techniques to set priorities for research will have to prove their value. It is possible that cost-effectiveness analysis may enhance its influence if it can address key methodological challenges in the measurement of benefits and costs, and techniques for sensitivity analysis. It is also possible that resistance to the use of cost-effectiveness analysis in decisions relating to policy, such as allocation of research funds, may be less than the resistance to its use in decisions concerning the allocation of medical treatments. Nevertheless, the use of formal techniques to help set priorities for research seems more likely to gain acceptance through instances where neglected areas of research can be identified as important through formal analysis than through instances where research is suggested to be of little value. Consistent with this, it appears that threats to planned increases in the NIH budget due to questions from some members of Congress about the value of increases in appropriations for research and concerns about NIH priorities in allocating research funds were an important motivation for the Institute of Medicine report that encouraged efforts to use formal approaches to determine the value of research.

References

Al, Maiwenn J., Ben A. van Hout, Bowine C. Michel, and Frans F.H. Rutten. "Sample Size Calculation in Economic Evaluations" Health Economics 7:327-35, 1998.

Arrow, Kenneth. Social Choice and Individual Values. Yale University Press, New Haven, CT, 1951.

Arrow, Kenneth and Robert Lind. "Uncertainty and the Evaluation of Public Investment Decisions" American Economic Review 16, 3:364-378, June 1970.

Briggs, Andrew H. and Alastair M. Gray. "Power and Significance Calculations for Stochastic Cost-Effectiveness Analysis" Medical Decision Making 18, supplement: S81-S92, 1998.

Brook, Robert, et al. "Does Free Care Improve Adults' Health? Results from a Randomized Controlled Trial," New England Journal of Medicine 309(24): 1426-1434, 1983.

Claxton, Karl. The Irrelevance of Inference: a Decision-Making Approach to the Stochastic Evaluation of Health Care Technologies." Forthcoming, Journal of Health Economics.

Claxton, Karl and John Posnett. "An Economic Approach to Clinical Trial Design and Research Priority-Setting," Health Economics 513-24, 1996.

Cutler, David and Elizabeth Richardson. "Measuring the Health of the U.S. Population," Brookings Papers: Microeconomics 1997: 217-271.

Eddy, David, 1990. "Screening for Cervical Cancer," Annals of Internal Medicine 113, 214-26.

Elixhauser, Anne, Michael Halpern, Jordana Schmier and Bryan R Luce, 1998. "Health Care CBA and CEA from 1991 to 1996: An Updated Bibliography". Medical Care 36 (5): MS1-MS9.

Felli, James C. and Gordon B. Hazen. "Sensitivity Analysis and the Expected Value of Perfect Information," Medical Decision Making 18 (1): 95-109, Jan-Mar 1998.

Fuchs, Victor. "The Health Sector's Share of the Gross National Product," Science 1990; 247: 534-538.

Gold, Marthe R., Joanna E. Siegel and Louise B. Russel, Milton C. Weinstein, 1996..Cost-Effectiveness in Health and Medicine. Oxford University Press, New York.

Hornberger, John. "A Cost-Benefit Analysis of a Cardiovascular Disease Prevention Trial, Using Folate Supplementation as an Example," American Journal of Public Health 88 (1): 61-67, 1998.

Institute of Medicine. Scientific Opportunities and Public Needs: Improving Priority Setting and Public Input at the National Institutes of Health. National Academy Press, Washington, DC, 1998.

Kahneman, Daniel and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." Econometrica 47(2):263-291, 1979.

McClellan, Mark, Barbara J. McNeil, and Joseph P. Newhouse. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?" JAMA 272(11): 859-866, September 21, 1994.

Meltzer, David. "Accounting for Future Costs in Medical Cost-Effectiveness Analysis". Journal of Health Economics. 16 (1): 33-64, 1997.

Meltzer, David and Magnus Johannesson. "Inconsistencies in the 'Societal Perspective' on Costs of the Panel on Cost-Effectiveness in Health and Medicine". Unpublished manuscript, 1998.

Meltzer, David, Tamar Polonsky, Brian Egelston and Janet Tobian. "Does Quality-adjusted Life Expectancy Predict Patient Preferences for Treatment in IDDM: Validating Quality of Life Measures through Revealed Preference." Unpublished manuscript, 1998.

Mullahy, John. "Which Cost-Effectiveness Ratio? Evaluating Health Policies and Medical Technologies in Stochastic Contexts". Unpublished manuscript, University of Wisconsin, Madison.

Murphy, Kevin M. and Robert H. Topel, "The Economic Value of Medical Research". Unpublished manuscript, University of Chicago, 1998.

O'Brien, Bernie J. Michael F. Drummond, Roberta J. Labelle, and Andrew Willian. "In Search of Power and Significance: Issues in the Design and Analysis of Stochastic Cost-Effectiveness Studies in Health Care," Medical Care 32 (2): 150-163, 1994.

Newhouse, Joseph P. "Medical Care Costs: How Much Welfare Loss." Journal of Economic Perspectives 6, 3-21, 1992.

Polsky, Daniel, Henry A Glick, Richard Willke and Kevin Shulman. "Confidence Intervals for Cost-Effectiveness Analysis: A Comparison of Four Methods" Health Economics 6:243-252, 1997.

Pratt, John W., Howard Raiffa, and Robert Schlaifer. Introduction to Statistical Decision Theory. McGraw-Hill Book Company, New York, 1965.

Raiffa, Howard and Robert Schlaifer. Applied Statistical Decision Theory. Harvard Business School, Colonial Press, 1961.

Rothschild, Michael and Joseph Stiglitz. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information" Quarterly Journal of Economics 90: 629-650, 1976.

Smith, Sheila et al. "The Next Ten Years of Health Spending: What Does the Future Hold?" Health Affairs 17 (5): 128-40, 1998.

Stinnett, Aaron A. and John Mullahy. "Net Health Benefits: A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis" Medical Decision Making 18(2 Suppl): S68-80, 1998.

Stinnett, Aaron A. and David Paltiel. "Estimating CE Ratios under Second-order Uncertainty: The Mean Ratio versus the Ratio of the Means," Medical Decision Making 17: 483-9, 1997.

Tambour, Magnus, Niklas Zethraeus, and Magnus Johannesson, International Journal of Technology Assessment in Health Care 14(3): 467-71, 1998

Table 1: Value of Information for Cost-Effectiveness of Screening for Prostate Cancer

Panel 1: Prostate Cancer Known Non-Aggressive: Fraction rapidly progressing = 0.085

Strategy	Cost (K\$)	ΔCost (K\$)	QALY	ΔQALY	ΔC/ΔQALY (K\$/QALY)	Value ΔQALY (K\$)*	Net Increm. Benefit (K\$)
Watchful Waiting	475	.	23.725
Radical Prostatectomy	485	11	23.75	0.025	420	2.5	-8.5

Panel 2: Prostate Cancer Known Aggressive: Fraction rapidly progressing = 0.2

Strategy	Cost (K\$)	ΔCost (K\$)	QALY	ΔQALY	ΔC/ΔQALY (K\$/QALY)	Value ΔQALY (K\$)*	Net Increm. Benefit (K\$)
Watchful Waiting	440	.	22
Radical Prostatectomy	485	45	23.75	1.75	26	175	130

Panel 3: Aggressiveness of Prostate Cancer not Known: Prob. Aggressive (as in Panel 2) = .2

Strategy	Cost (K\$)	ΔCost (K\$)	QALY	ΔQALY	ΔC/ΔQALY (K\$/QALY)	Value ΔQALY (K\$)*	Net Increm. Benefit (K\$)
Watchful Waiting	467.6	.	23.38
Radical Prostatectomy	485	17.4	23.75	0.37	47	37	19.6

Panel 4: Expected Value with Perfect Information: Prob. Aggressive (as in Panel 2) = 0.2

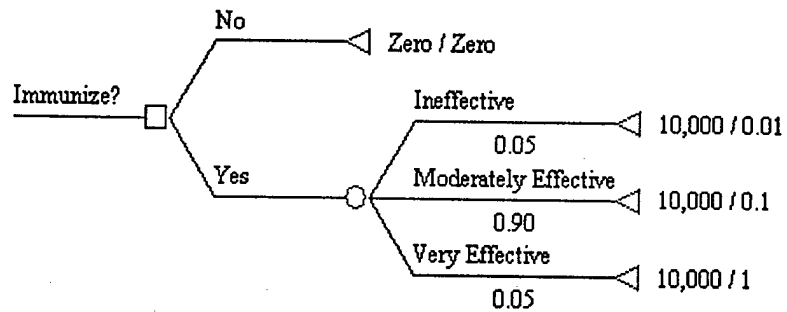
Strategy	Cost (K\$)	ΔCost (K\$)	QALY	ΔQALY	ΔC/ΔQALY (K\$/QALY)	Value ΔQALY (K\$)*	Net Increm. Benefit (K\$)
Watchful Waiting	467.6	.	23.38
Optimal w/ Perfect Info.	476.6	9	23.73	0.35	26	35	26
Radical Prostatectomy	485	8.4	23.75	0.02	420	2	-6.4

Panel 5: Expected Value w/ Improved Info.: 50% chance study \Rightarrow Prob. Agg.=0.05; 50% chance study \Rightarrow Prob. Agg. =0.35

Strategy	Cost (K\$)	ΔCost (K\$)	QALY	ΔQALY	ΔC/ΔQALY (K\$/QALY)	Value ΔQALY (K\$)*	Net Increm. Benefit (K\$)
Prob. Aggressive=0.05							
Watchful Waiting	473.3	.	23.64
Radical Prostatectomy	485	11.7	23.75	0.11	106	11.1	-0.6
Prob. Aggressive=0.35							
Watchful Waiting	462.8	.	23.12
Radical Prostatectomy	485	22.2	23.75	0.63	35	62.9	40.6
Expected Value of Info. (vs. expected value of optimal decision with initial info. (Prostatectomy)) = $0.5 \times 0.6 =$							0.3

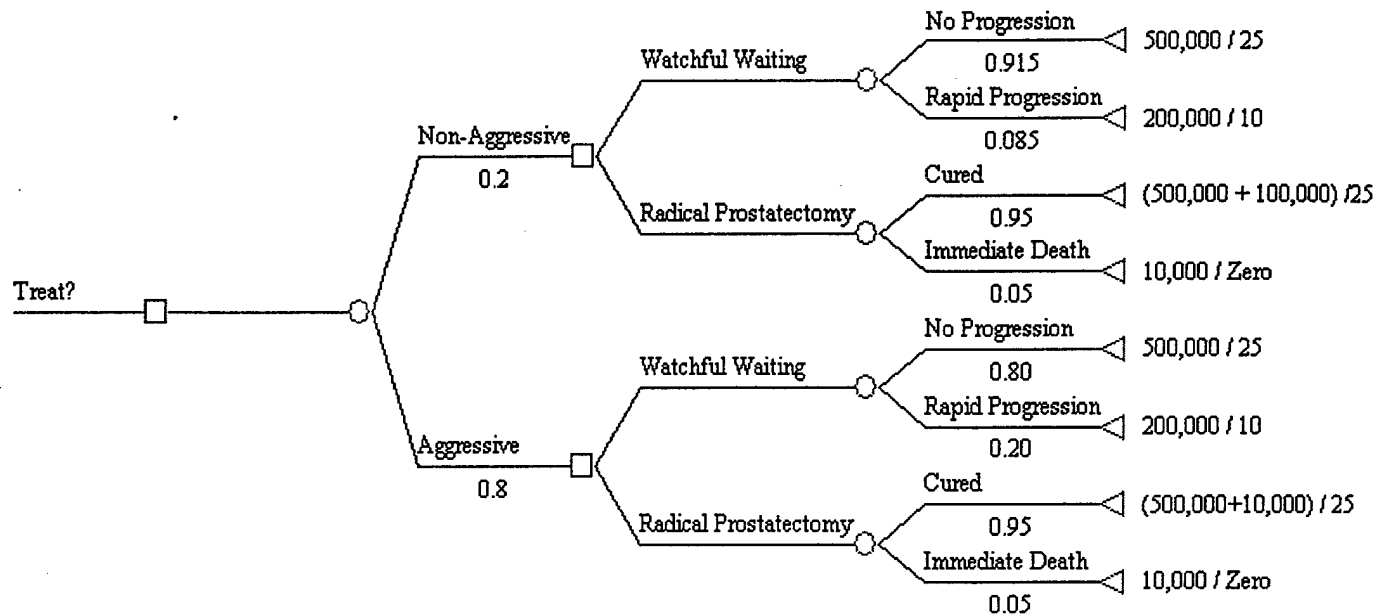
* 1QALY=\$100K

Figure 1: Simplified Decision Concerning a Treatment of Uncertain Benefit



Legend: Cost (\$) / Effectiveness (Life Years)

Figure 2: Simplified Cost-Effectiveness Model for Screening for Prostate Cancer with
Uncertainty about Progression Rates



Legend: Cost (\$) / Effectiveness (Life Years)